

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 07:54:35 ; Search time 3376 Seconds

(without alignments)
654.767 Million cell updates/sec

Title: US-10-032-393-36

Perfect score: 51
Sequence: 1 tcataaaattttttgtct.....tttctgtataatagattca 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	100.0	60	6	A15398	A15398 Synthetic P
2	51	100.0	75	7	PT5PN25	M11602 Bacterioph
3	51	100.0	202	6	A10374	A10374 Nucleotide
4	51	100.0	202	6	A11246	A11246 vector pDS2
5	51	100.0	202	6	E00936	E00936 DNA sequenc
6	51	100.0	207	6	A10375	A10375 Nucleotide
7	51	100.0	247	6	A11242	A11242 vector pDS1
8	51	100.0	259	6	A10373	A10373 Nucleotide
9	51	100.0	1113	6	A02739	A02739 Artificial
10	51	100.0	1113	6	A14592	A14592 Synthetic n
11	51	100.0	1221	6	A02296	A02296 Plasmid pDS
12	51	100.0	1221	6	A14593	A14593 Synthetic n
13	51	100.0	1866	6	A02227	A02227 Plasmid pDS
14	51	100.0	1866	6	A02295	A02295 Plasmid pDS
15	51	100.0	1866	6	A02740	A02740 Artificial
16	51	100.0	1866	6	A07053	A07053 Nucleotide
17	51	100.0	1866	6	A14594	A14594 Synthetic n
18	51	100.0	3403	6	A02228	A02228 Plasmid pDS
19	51	100.0	5767	6	AX035965	AX035965 Sequence
20	51	100.0	6000	12	U66308	U66308 Expression
21	51	100.0	6447	12	AF288421	AF288421 Synthetic
22	50	98.0	76	6	A11173	A11173 A TS promot
23	46.4	91.0	524	12	SYNIFNGS	K01699 Human immu
24	46	90.2	73	6	A12013	A12013 oligonucleo
25	46	90.2	73	6	A12014	A12014 oligonucleo
26	42.8	83.9	83	6	A16630	A16630 Nucleotide
27	42.8	83.9	83	6	A16631	A16631 Nucleotide
28	39	76.5	68	6	E01976	E01976 DNA encodin
29	33.6	65.9	131	6	A50146	A50146 Sequence 3
30	33.4	65.5	125	6	A46760	A46760 Sequence 2
31	33.4	65.5	125	6	I89341	I89341 Sequence 2
32	33	64.7	116	6	A13088	A13088 Nucleotide
33	33	64.7	132	12	SYNADRE1B1	M1661 Synthetic a
34	33	64.7	171	6	AR069843	AR069843 Sequence
35	33	64.7	173	6	A01257	A01257 Nucleotide
36	33	64.7	173	6	A01258	A01258 (reverse co
37	33	64.7	173	6	A13104	A13104 Nucleotide
38	33	64.7	173	6	A13105	A13105 Nucleotide
39	33	64.7	180	6	A01247	A01247 Nucleotide
40	33	64.7	180	6	A01248	A01248 (reverse co
41	33	64.7	180	6	A13094	A13094 Nucleotide
42	33	64.7	180	6	A13095	A13095 Nucleotide
43	33	64.7	319	12	ARLACT	X06035 Synthetic 1
44	33	64.7	396	6	A11819	A11819 Artificial
45	33	64.7	396	6	A11820	A11820 Artificial

ALIGNMENTS

RESULT 1
A15398
LOCUS A15398 60 bp DNA linear PAT 22-SEP-1995
DEFINITION Synthetic Phage T5 promoter P-N 25.
ACCESSION A15398
VERSION A15398.1 GI:1247805
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 60)
AUTHORS Bujard, H. and Le Grice, S.
TITLE New gram-positive expression-control sequences
JOURNAL Patent EP 0207459-A2 07-JAN-1987;
F. HOFFMANN-VA ROCHÉ AG

FEATURES
Location/Qualifiers
1..60

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51

Db 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51

RESULT 2
PTSPN25

LOCUS PTSPN25 75 bp DNA linear PHG 28-APR-1993
DEFINITION Bacteriophage T5 promoter P-N 25.
ACCESSION M11602
VERSION M11602.1 GI:215985
KEYWORDS promoter region.
SOURCE Bacteriophage T5
ORGANISM Bacteriophage T5

Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
T5-like viruses.

1 (bases 1 to 75)

Gentz,R. and Bujard,H.

Promoters recognized by Escherichia coli RNA polymerase selected by
function: highly efficient promoters from Bacteriophage T5

JOURNAL J. Bacteriol. 164 (1), 70-77 (1985)

MEDLINE 86008105

PUBMED 3900050

COMMENT Original source text: Bacteriophage T5 DNA.

FEATURES
source

1..75

/organism="Bacteriophage T5"

/mol_type="genomic DNA"

/db_xref="taxon:10726"

56..>75

/note="bacteriophage T5 RNA"

ORIGIN
Undetermined.

Query Match 100.0%; Score 51; DB 7; Length 75;

Best Local Similarity 100.0%; Pred. No. 0.0085;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51

Db 6 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 56

RESULT 3

LOCUS A10374

DEFINITION Nucleotide sequence 21 from patent number EP0303925.
ACCESSION A10374
VERSION A10374.1 GI:490696

KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

1 (bases 1 to 202)

Bujard,H. and Lanzer,M.

High repressible sequence for control of expression
Patent: EP 0303925-A 21 22-FEB-1989;

F. HOFFMANN-LA ROCHE AG

FEATURES
Location/Qualifiers
1..202

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 202;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51

Db 124 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 174

RESULT 4

LOCUS A11246

DEFINITION vector pDS2/PN25x/0, to2+ XhoI/EcoRI fragment.
ACCESSION A11246
VERSION A11246.1 GI:491020

KEYWORDS

synthetic construct
synthetic construct
artificial sequences.

1 (bases 1 to 202)

Bujard,H. and Stueber,D.

New expression control sequence
Patent: EP 0186069-A 10 02-JUL-1986;

F. HOFFMANN-LA ROCHE AG

FEATURES
Location/Qualifiers
1..202

source

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 202;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51

Db 124 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 174

RESULT 5

LOCUS E00936

DEFINITION DNA sequence including promoter and operator hybrid 'P(N25X/)'.
ACCESSION E00936
VERSION E00936.1 GI:2169197

KEYWORDS JP 1986181386-A/1.

Escherichia coli

Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

1 (bases 1 to 202)

Heruman,B. and Deietoritsuhi,S.

NOVEL DEVELOPMENT REGULATING ARRANGEMENT
Patent: JP 1986181386-A 1 14-AUG-1986;

F HOFFMANN LA ROCHE & CO AG

OS Escherichia coli

PN JP 1986181386-A/1

PD 14-AUG-1986

PF 16-DEC-1985 JP 1985282699

PR 17-DEC-1984 GB 84 8431818

PI HERUMAN BUYARUDO, DEIETORITSUHI SUTHEYUBAA

PC C12N15/00, C12N1/20, C12N9/10, C12P21/02, (C12N1/20, C12R1:19), PC

(C12N1/20, C12R1:125), (C12N9/10, C12R1:19), (C12N9/10, C12R1:125), (C12P21/02, PC

C12R1:19),

PC (C12P21/02, C12R1:125);

CC strandedness: Single;

CC topology: Linear;

CC hypothetical: No;

```
CC anti-sense: No;
CC *source: strain=M15;
FH Key Location/Qualifiers
FT promoter 1..<173
FT /note='P25 promoter'
FT sig_peptide 174..>190
FT /note='lac operator'

FEATURES
    source
        Location/Qualifiers
            1..202
                /organism="Escherichia coli"
                /mol_type="genomic DNA"
                /db_xref="taxon:562"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 202;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
    124 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 174

Db 124 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 174

RESULT 6
LOCUS A10375 207 bp DNA linear PAT 22-SEP-1993
DEFINITION Nucleotide sequence 22 from patent number EP0303925.
ACCESSION A10375
VERSION A10375.1 GI:490697
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 207)
AUTHORS Bujard,H. and Lanzer,M.
TITLE High repressible sequence for control of expression
JOURNAL Patent: EP 0303925-A 22 22-FEB-1989;
          F. HOFFMANN-LA ROCHE AG

FEATURES
    source
        Location/Qualifiers
            1..207
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
    124 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 174

Db 124 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 174

RESULT 7
LOCUS A11242 247 bp DNA linear PAT 12-NOV-1993
DEFINITION vector pDS1/PN25 XhoI-fragment carrying promoter Pn25 is displayed.
ACCESSION A11242
VERSION A11242.1 GI:491016
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 247)
AUTHORS Bujard,H. and Stueber,D.
TITLE New expression control sequence
JOURNAL Patent: EP 0186069-A 6 02-JUL-1986;
          F. HOFFMANN-LA ROCHE AG

FEATURES
    source
        Location/Qualifiers
            1..247
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /translation="MRGSVDLQPSLARFSGAKKMEKKITGYTTVDISQMRKEHFE"
                /organism="synthetic construct"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
    124 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 174

Db 124 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 174

RESULT 8
LOCUS A10373 259 bp DNA linear PAT 22-SEP-1993
DEFINITION Nucleotide sequence 20 from patent number EP0303925.
ACCESSION A10373
VERSION A10373.1 GI:490695
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 259)
AUTHORS Bujard,H. and Lanzer,M.
TITLE High repressible sequence for control of expression
JOURNAL Patent: EP 0303925-A 20 22-FEB-1989;
          F. HOFFMANN-LA ROCHE AG

FEATURES
    source
        Location/Qualifiers
            1..259
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
    130 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 180

Db 130 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 180

RESULT 9
LOCUS A02739 1113 bp DNA linear PAT 23-MAR-1993
DEFINITION Artificial sequence of plasmid pDS5/RBSII,3A+5A (XhoI/XbaI
          fragment).
ACCESSION A02739
VERSION A02739.1 GI:345272
KEYWORDS chloramphenicol acetyltransferase.
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 1113)
AUTHORS Gentz,R., Le Grice,S., Mous,J. and Stueber,D.
TITLE ENV/GAG polypeptides
JOURNAL Patent: EP 0270114-A 5 08-JUN-1988;
          F. HOFFMANN-LA ROCHE AG

FEATURES
    source
        Location/Qualifiers
            1..1113
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                225..947
                    /codon_start=1
                    /transl_table=11
                    /product="chloramphenicol acetyltransferase"
                    /protein_id="CAA00265.1"
                    /db_xref="GI:345273"
                    /db_xref="REMBL:CAA00265"
                    /translation="MRGSVDLQPSLARFSGAKKMEKKITGYTTVDISQMRKEHFE"
                    /organism="synthetic construct"
```

AFQSVAAQCTYNQTVQLDITAE LKTVKKNKHKFPAPFIHILARLMAHPERPMAMKDG
LVIMDSVHPCYTVFHEQETEFSSIMSEYHDDFRQFLHYSQDVACYGENLAYFPKGF
ENMFVVSANPMWVSFTSFDLNVAANMDNFPAPVFTMGKYYTQGDKVLMPLAIOVHHA
VCD GFHVGRMLNELQOYCDWQGA"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1113;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

RESULT 10

LOCUS A14592 1113 bp DNA circular PAT 21-FEB-1994

DEFINITION Synthetic nucleotide sequence encoding chloramphenicol

acetyltransferase (plus 21 additional N-terminus amino acids).

ACCESSION A14592

VERSION A14592.1 GI:491830

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 1113)

AUTHORS Banwarth,W., Certa,U., Mous,J. and Stueber,D.

TITLE Polypeptides that elicit antibodies against AIDS virus

JOURNAL Patent: EP 0219106-A 5 22-APR-1987;

F. HOFFMANN-LA ROCHE AG

FEATURES Location/Qualifiers

source 1..1113

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

225..947

/codon_start=1

/transl_table=11

/product="chloramphenicol acetyltransferase"

/protein_id="CAA01171.1"

/db_xref="GI:491831"

/db_xref="REMTREMBL:CAA01171"

/translation="MRGSVDLPISLARFSGAKKMKKITGYTTVDISQMRKHEFE

AFQSVAAQCTYNQTVQLDITAE LKTVKKNKHKFPAPFIHILARLMAHPERPMAMKDG

LVIMDSVHPCYTVFHEQETEFSSIMSEYHDDFRQFLHYSQDVACYGENLAYFPKGF

ENMFVVSANPMWVSFTSFDLNVAANMDNFPAPVFTMGKYYTQGDKVLMPLAIOVHHA

VCD GFHVGRMLNELQOYCDWQGA"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1113;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

RESULT 11

LOCUS A02296 1221 bp DNA linear PAT 18-MAY-1993

DEFINITION Plasmid pDS6/RBSII, 3A+5A DNA XhoI/XbaI-fragment.

ACCESSION A02296

VERSION A02296.1 GI:345259

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 1221)

AUTHORS Certa,U., Gentz,R. and Takacs,B.

TITLE Plasmodium falciptarum merozoite antigen peptides

JOURNAL Patent: EP 0283829-A 17 28-SEP-1988;

F. HOFFMANN-LA ROCHE AG
Location/Qualifiers
1..1221

FEATURES

source /organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1221;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

RESULT 12

LOCUS A14593 1221 bp DNA circular PAT 21-FEB-1994

DEFINITION Synthetic nucleotide sequence of the XhoI/XbaI fragment of

pDS6/RBSII, 3A+5A.

ACCESSION A14593

VERSION A14593.1 GI:491832

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 1221)

AUTHORS Banwarth,W., Certa,U., Mous,J. and Stueber,D.

TITLE Polypeptides that elicit antibodies against AIDS virus

JOURNAL Patent: EP 0219106-A 6 22-APR-1987;

F. HOFFMANN-LA ROCHE AG

FEATURES Location/Qualifiers

source 1..1221

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

225..269

/gene="cat"

/gene="cat"

/codon_start=1

/transl_table=11

/protein_id="CAA01172.1"

/db_xref="GI:491833"

/db_xref="REMTREMBL:CAA01172"

/translation="MRGSVDLPISLSDSC"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1221;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

RESULT 13

LOCUS A02227 1866 bp DNA linear PAT 29-APR-1996

DEFINITION Plasmid pDS8/RBSII, SphI DNA for XhoI/XbaI fragment.

ACCESSION A02227

VERSION A02227.1 GI:490286

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 1866)

AUTHORS Doebeili,H., Eggimann,B., Gentz,R., Hochuli,E. and Stueber,D.

TITLE Fusion proteins and their purification

JOURNAL Patent: EP 0282042-A 1 14-SEP-1988;

FEATURES
source
F. HOFFMANN-LA ROCHE AG
Location/Qualifiers
1..1866

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1866;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

RESULT 14
A02295 1866 bp DNA linear PAT 18-MAY-1993

LOCUS A02295 Plasmid pDS8/RBSII, SphI DNA XhoI/XbaI-fragment.
DEFINITION A02295
ACCESSION A02295
VERSION A02295.1 GI:345258

KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1 (bases 1 to 1866)
Certa, U., Gentz, R. and Takacs, B.

AUTHORS Plasmodium falciparum merozoite antigen peptides
TITLE Patent: EP 0283829-A 16 28-SEP-1988;
JOURNAL F. HOFFMANN-LA ROCHE AG

FEATURES
source Location/Qualifiers
1..1866
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1866;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

RESULT 15
A02740 1866 bp DNA linear PAT 23-MAR-1993

LOCUS A02740 Artificial sequence of pDS8/RBSII plasmid (XhoI/XbaI fragment).
DEFINITION A02740
ACCESSION A02740
VERSION A02740.1 GI:345274

KEYWORDS dihydrofolate reductase.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1 (bases 1 to 1866)
Gentz, R., Le Grice, S., Mous, J. and Stueber, D.

AUTHORS ENV/GAG polypeptides
TITLE Patent: EP 0270114-A 6 08-JUN-1988;
JOURNAL F. HOFFMANN-LA ROCHE AG

FEATURES
source Location/Qualifiers
1..1866
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
225..806
/codon_start=1
/transl_table=11
/product="dihydrofolatreductase"
/protein_id="CAA00266.1"

/db_xref="GI:345275"
/db_xref="REMBL:CAA00266"
/translation="MRSGIMVRPLNCIVAVSQMGIGKNGDLPWPLRNEFKYFQRM
TTSSVEGKQNLVIMGRKTWFSIPKQNRPLKORINIVLSRELKEPRGAFLAKSLDD
ALRLIEQPELASKVDWVWIVGSSVYQEAMNQPGLRLFTVTRINQEFESDTFPPEIDL
GKYKLPRYPGVLSVQEEKGIKYKFEYEEKD"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1866;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
125 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 175

Search completed: May 4, 2004, 11:34:14
Job time : 3379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 07:51:37 ; Search time 389 Seconds
(without alignments)
556.962 Million cell updates/sec

Title: US-10-032-393-36
Perfect score: 51
Sequence: 1 tcataaaaaatttattgtct.....ttttctgtataatagattca 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1908s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	51	6	ABK98620 Gram posi
2	51	100.0	51	8	ACD13871 Xyl-T5 pr
3	51	100.0	60	1	AAN70573 Coliphage
4	51	100.0	72	4	AAH74868 Oligonuc1
5	51	100.0	80	4	AAH74869 Oligonuc1
6	51	100.0	80	6	ABK98610 Gram posi
7	51	100.0	80	8	ACD13861 Xyl-T5 fu
8	51	100.0	94	6	ABK98589 Xy10-T5 f
9	51	100.0	94	6	ACD13840 Oligonuc1
10	51	100.0	95	6	ABK98590 Xy10-T5 f
11	51	100.0	95	8	ACD13841 Oligonuc1
12	51	100.0	200	1	AAN91066 Control s
13	51	100.0	202	1	AAN60262 Plasmid p
14	51	100.0	207	1	AAN91067 Control s
15	51	100.0	247	1	AAN60259 Coliphage
16	51	100.0	259	1	AAN91065 Control s
17	51	100.0	556	6	ABK98585 Gram posi
18	51	100.0	556	8	ACD13836 Xyl-T5 fu
19	51	100.0	1113	1	AAN80481 Xhol/Xba1
20	51	100.0	1221	1	AAN81153 Xho1/Xba1
21	51	100.0	1866	1	AAN80482 Xho1/Xba1
22	51	100.0	1866	1	AAN81356 Sequence
23	51	100.0	1866	1	AAN81152 Xho1/Xba1

24	51	100.0	1866	1	AAN80955 Xho1-Xba1
25	51	100.0	3403	1	AAN80956 Plasmid p
26	51	100.0	5302	3	AAA98012 Expressio
27	51	100.0	5767	3	AAA98014 Expressio
28	51	100.0	6852	6	ABK98600 Vector pX
29	51	100.0	6852	8	ACD13851 Xyl opera
30	50	98.0	1246	4	AAS50517 Staphyloc
31	46	90.2	77	1	AAN70113 Sequence
32	43	84.3	51	6	ABK98621 Gram posi
33	43	84.3	51	8	ACD13872 Xyl-T5 mu
34	43	84.3	80	6	ABK98611 Gram posi
35	43	84.3	80	6	ACD13862 Xyl-T5 mu
36	43	84.3	94	6	ABK98606 Xyl-T5 DD
37	43	84.3	94	8	ACD13857 Xyl-T5 mu
38	43	84.3	95	6	ABK98607 Xyl-TS DD
39	43	84.3	6852	6	ABK98637 Vector pX
40	43	84.3	6852	8	ACD13888 Xyl-T5 mu
41	42.8	83.9	83	1	AAN81195 Beta-lact
42	39.4	77.3	80	1	AAN80462 Expressio
43	39	76.5	68	1	AAN81260 Hybrid pr
44	36.4	71.4	41	1	AAN90998 Fragments
45	33.4	65.5	125	2	AAT04945 Fragment

ALIGNMENTS

RESULT 1	
ABK98620	
ID	ABK98620 standard; DNA; 51 BP.
XX	
AC	ABK98620;
XX	
DT	07-AUG-2003 (revised)
DT	21-OCT-2002 (first entry)
XX	
DE	Gram positive bacteria Xyl-T5 fusion promoter #3.
XX	
KW	ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;
KW	P59; P1P2; PU; xy10; teto; trpO; malo; lambdaclO; cellular proliferation;
KW	antibiotic.
XX	
OS	Eubacteria.
OS	Bacteriophage T5.
OS	Synthetic.
XX	
PN	WO200251982-A2
XX	
PD	04-JUL-2002.
XX	
PF	21-DEC-2001; 2001WO-US050250.
XX	
PR	27-DEC-2000; 2000US-0259434P.
PR	06-SEP-2001; 2001US-00948993.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
PI	Haselbeck R, Wall D, Gross M;
XX	
DR	WPI; 2002-575374/61.
XX	
PT	Isolated nucleic acid comprises bacterial promoters modified to have
PT	altered activity in at least one gram-positive organism, e.g. Bacillus
PT	anthracis or Clostridium botulinum, useful for regulating gene expression
PT	in bacteria.
XX	
PS	Claim 2; Page 219; 246pp; English.
XX	
CC	The invention relates to an isolated nucleic acid comprising a fusion
CC	promoter comprising at least one promoter that is modified to have
CC	altered activity in at least one gram-positive organism, or comprising
CC	T5, CP25, P32, P59, P1P2 or PU linked to at least one operator consisting
CC	of xy10, teto, trpO, malo or lambdaclO, where at least one operator is

CC positioned so binding of a repressor to an operator represses
CC transcription from the fusion promoter. Also included are vectors and
CC host cells comprising the fusion promoters, a method of identifying genes
CC involved in cellular proliferation or required for proliferation of a
CC prokaryotic cell using the vector, a method of identifying compounds that
CC inhibit the proliferation of a prokaryotic cell using the vector, a
CC method of identifying a compound that reduces the activity or level of a
CC gene product required for proliferation of a cell using the vector, a
CC compound identified by the methods, a method of inhibiting the activity
CC or expression of a gene in an operon required for proliferation using the
CC vector, manufacturing an antibiotic comprising using the vector or cell
CC and identifying a nucleic acid with promoter activity in *Enterococcus*
CC faecalis. The fusion promoters are useful for regulating nucleic acid or
CC polypeptide expression, particularly for regulating gene expression in
CC bacteria and for identifying proliferation-regulated genes or molecules
CC with potential antibiotic activity. The modified promoters are also
CC useful for replacing endogenous promoters to create cells with specific
CC regulatable genes. The present sequence is a fusion promoter sequence of
CC the invention. (Updated on 07-AUG-2003 to correct OS field.)

XX SQ Sequence 51 BP; 19 A; 5 C; 5 G; 22 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
Db 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 2

ACD13871
ID ACD13871 standard; DNA; 51 BP.

XX AC ACD13871;

XX DT 15-AUG-2003 (first entry)

XX DE Xyl-T5 promoter sequence.

KM Promoter; ds; gram positive bacteria; *Staphylococcus aureus*;
KM *Enterococcus faecalis*; operator; xylO; tetO; trpO; malO; lambda-clO;
KM cellular proliferation.

OS Bacteriophage T5.
OS Unidentified.

XX PN US2003027286-A1.

XX PD 06-FEB-2003.

XX PF 21-DEC-2001; 2001US-00032393.

XX PR 06-SEP-2000; 2000US-0230335P.

XX PR 27-DEC-2000; 2000US-0259434P.

XX PA (HASE/) HASELBECK R.

PA (WALL/) WALL D.

PA (GROS/) GROSS M.

XX PI Haselbeck R, Wall D, Gross M;

XX DR WPI; 2003-479541/45.

XX PT New isolated nucleic acid comprising a fusion promoter having at least
PT one promoter that is modified to have altered activity in at least one
PT gram-positive organism, useful for regulating gene expression in
PT bacteria.

PS Claim 2; Page 80; 142pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion

CC promoter having at least one promoter that is modified to have altered
CC activity in at least one gram-positive organism (e.g. *Staphylococcus*
CC *aureus* or *Enterococcus faecalis*). The promoter is linked to at least one
CC operator selected from xylO, tetO, trpO, malO and lambda-clO, which are
CC positioned such that the binding of at least one repressor to the
CC operator represses transcription from the fusion promoter. Also included
CC are a vector comprising the isolated nucleic acid, a host cell comprising
CC the nucleic acid. The fusion promoter is useful for identifying genes
CC involved in cellular proliferation, identifying a compound that reduces
CC the activity or level of a gene product required for proliferation of a
CC cell, inhibiting the activity or expression of a gene in an operon
CC required for proliferation, manufacturing an antibiotic, identifying a
CC gene that is required for proliferation of a prokaryotic cell,
CC identifying a compound that inhibits the proliferation of a prokaryotic
CC cell and regulating gene expression in bacteria. The present sequence is
CC a bacterial promoter suitable for inclusion in a fusion promoter of the
CC invention

XX SQ Sequence 51 BP; 19 A; 5 C; 5 G; 22 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 8; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
Db 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 3

AAN70573
ID AAN70573 standard; DNA; 60 BP.

XX AC AAN70573;

XX DT 25-MAR-2003 (revised)

XX DT 11-MAR-1991 (first entry)

XX DE Coliphage T5 promoter of *Bacillus subtilis* expression vector.

KM Expression vector; dihydrofolate reductase; IL-2; insulin; HTLV-III; tPA;
KM human renin; ds.

XX OS Coliphage T5.

XX FH Key Location/Qualifiers

FT -35_signal 16..21

FT FT /*tag= a

FT -10_signal 39..44

FT FT /*tag= b

XX PN BP207459-A.

XX PD 07-JAN-1987.

XX PF 27-JUN-1986; 86BP-00108774.

XX PR 05-JUL-1985; 85GB-00017071.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Bujard H, Legrice S;

XX DR WPI; 1987-001183/01.

XX PT New gram positive expression control DNA sequences - useful for potent
PT and versatile gene expression of prokaryotic or eukaryotic proteins in
PT *Bacillus subtilis* etc.

PS Disclosure; Table I; 53pp; English.

XX The gene fragment may be incorporated into a novel expression vector, for
CC controlling protein production in transformed Gram-positive bacteria esp.

CC Bacillus subtilis. Proteins which may be encoded include: dihydrofolate
CC reductase; chloramphenicol acetyltransferase; malaria surface antigens;
CC IL-2; interferons; insulin; tPA; human renin and HTLV-III. (Updated on 25
CC -MAR-2003 to correct PI field.)

XX
SQ Sequence 60 BP; 23 A; 5 C; 6 G; 26 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
1 TCATAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
Db

RESULT 4

AAH74868
ID AAH74868 standard; DNA; 72 BP.

XX
AC AAH74868;

XX
DT 29-OCT-2001 (first entry)

DE Oligonucleotide used to create T5 promoter.

XX
KW Recombinase; genetic engineering; monogenic disorder; ADA deficiency;
KW cystic fibrosis; familial-hypercholesterolaemia; haemophilia;
KW chronic granulomatous disease; Duchenne's muscular dystrophy;
KW Fanconi's disease; anemia; sickle-cell anemia; Gaucher's disease;
KW Hunter's syndrome; X-linked severe combined immunodeficiency; SCID;
KW infectious disease; acquired disorder; tumour; cancer; T5 promoter; ss.
XX
OS Synthetic.

XX
PN WO200161049-A1.

XX
PD 23-AUG-2001.

XX
PF 16-FEB-2001; 2001WO-US005269.

XX
PR 18-FEB-2000; 2000US-0183759P.

XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX
PI Calos MP, Scilimenti CR;

XX
DR WPI; 2001-522610/57.

XX
PT Identifying altered recombinase, involves hybridizing recombination sites
PT in cells having specific coding sequence, transformed with altered
PT recombinase gene, and isolating cells having product of the sequence.

XX
PS Disclosure; Page 59; 101pp; English.

XX
CC The specification describes a method for identifying altered
CC recombinases. The method comprises transforming cells with a first
CC plasmid comprising two recombination sites and a coding sequence of
CC interest, and a second plasmid encoding an altered recombinase; allowing
CC recombination of the two recombination sites using the altered
CC recombinase; screening and isolating transformed cells comprising the
CC product of the sequence of interest; and identifying the altered
CC recombinase. The altered recombinase is useful for site-specifically
CC integrating a polynucleotide sequence of interest in a genome of a cell.
CC It is also useful in genetic engineering of chromosomes of higher cells,
CC and for the generation of transgenic cells, tissues, plants and animals.
CC The altered recombinase is useful for treating monogenic disorders, e.g.
CC ADA deficiency, cystic fibrosis, familial-hypercholesterolaemia, anemia,
CC chronic granulomatous disease, Duchenne's muscular dystrophy,
CC haemophilia, Fanconi's disease, sickle-cell anemia, Gaucher's disease,
CC Hunter's syndrome and X-linked severe combined immunodeficiency (SCID),
CC infectious diseases including viral and bacterial infections, acquired
CC disorders including solid tumours and haematopoietic cancers such as

CC leukaemias and lymphomas, and other cancers. Oligonucleotides AAH74868-69
CC were used to create the T5 promoter, which was used to construct a
CC plasmid for use in the course of the invention

XX
SQ Sequence 72 BP; 27 A; 6 C; 11 G; 28 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 4; Length 72;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
2 TCATAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 52
Db

RESULT 5

AAH74869/c
ID AAH74869 standard; DNA; 80 BP.

XX
AC AAH74869;

XX
DT 29-OCT-2001 (first entry)

DE Oligonucleotide used to create T5 promoter.

XX
KW Recombinase; genetic engineering; monogenic disorder; ADA deficiency;
KW cystic fibrosis; familial-hypercholesterolaemia; haemophilia;
KW chronic granulomatous disease; Duchenne's muscular dystrophy;
KW Fanconi's disease; anemia; sickle-cell anemia; Gaucher's disease;
KW Hunter's syndrome; X-linked severe combined immunodeficiency; SCID;
KW infectious disease; acquired disorder; tumour; cancer; T5 promoter; ss.
XX
OS Synthetic.

XX
PN WO200161049-A1.

XX
PD 23-AUG-2001.

XX
PF 16-FEB-2001; 2001WO-US005269.

XX
PR 18-FEB-2000; 2000US-0183759P.

XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX
PI Calos MP, Scilimenti CR;

XX
DR WPI; 2001-522610/57.

XX
PT Identifying altered recombinase, involves hybridizing recombination sites
PT in cells having specific coding sequence, transformed with altered
PT recombinase gene, and isolating cells having product of the sequence.

XX
PS Disclosure; Page 59; 101pp; English.

XX
CC The specification describes a method for identifying altered
CC recombinases. The method comprises transforming cells with a first
CC plasmid comprising two recombination sites and a coding sequence of
CC interest, and a second plasmid encoding an altered recombinase; allowing
CC recombination of the two recombination sites using the altered
CC recombinase; screening and isolating transformed cells comprising the
CC product of the sequence of interest; and identifying the altered
CC recombinase. The altered recombinase is useful for site-specifically
CC integrating a polynucleotide sequence of interest in a genome of a cell.
CC It is also useful in genetic engineering of chromosomes of higher cells,
CC and for the generation of transgenic cells, tissues, plants and animals.
CC The altered recombinase is useful for treating monogenic disorders, e.g.
CC ADA deficiency, cystic fibrosis, familial-hypercholesterolaemia, anemia,
CC chronic granulomatous disease, Duchenne's muscular dystrophy,
CC haemophilia, Fanconi's disease, sickle-cell anemia, Gaucher's disease,
CC Hunter's syndrome and X-linked severe combined immunodeficiency (SCID),
CC infectious diseases including viral and bacterial infections, acquired
CC disorders including solid tumours and haematopoietic cancers such as
CC leukaemias and lymphomas, and other cancers. Oligonucleotides AAH74868-69

CC were used to create the T5 promoter, which was used to construct a
CC plasmid for use in the course of the invention
XX
SQ Sequence 80 BP; 29 A; 14 C; 9 G; 28 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
|||||
75 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATATAGATTCA 25

RESULT 6
ABK98610
ID ABK98610 standard; DNA; 80 BP.

XX
AC ABK98610;
XX
DT 07-AUG-2003 (revised)
DT 21-OCT-2002 (first entry)

DE Gram positive bacteria Xyl-T5 fusion promoter #2.

XX
KM ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;
KM P59; P1P2; PL; xylO; teto; trpO; malO; lambdaclO; cellular proliferation;
KW antibiotic.

XX
OS Eubacteria.
OS Bacteriophage T5.
OS Synthetic.

XX
PN WQ200251982-A2.

XX
PD 04-JUL-2002.

XX
PF 21-DEC-2001; 2001WO-US050250.

XX
PR 27-DEC-2000; 2000US-0259434P.
PR 06-SEP-2001; 2001US-00948993.

XX
PA (ELIT-) ELITRA PHARM INC.

XX
PI Haselbeck R, Wall D, Gross M;

XX
DR WPI; 2002-575374/61.

XX
PT Isolated nucleic acid comprises bacterial promoters modified to have
PT altered activity in at least one gram-positive organism, e.g. *Bacillus*
PT anthracis or *Clostridium botulinum*, useful for regulating gene expression
PT in bacteria.

XX
PS Claim 24; Page 217; 246pp; English.

XX
CC The invention relates to an isolated nucleic acid comprising a fusion
CC promoter comprising at least one promoter that is modified to have
CC altered activity in at least one gram-positive organism, or comprising
CC T5, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting
CC of xylO, teto, trpO, malO or lambdaclO, where at least one operator is
CC positioned so binding of a repressor to an operator represses
CC transcription from the fusion promoter. Also included are vectors and
CC host cells comprising the fusion promoters, a method of identifying genes
CC involved in cellular proliferation or required for proliferation of a
CC prokaryotic cell using the vector, a method of identifying compounds that
CC inhibit the proliferation of a prokaryotic cell using the vector, a
CC method of identifying a compound that reduces the activity or level of a
CC gene product required for proliferation of a cell using the vector, a
CC compound identified by the methods, a method of inhibiting the activity
CC or expression of a gene in an operon comprising for proliferation using the
CC vector, manufacturing an antibiotic comprising using the vector or cell
CC and identifying a nucleic acid with promoter activity in *Enterococcus*
CC faecalis. The fusion promoters are useful for regulating nucleic acid or

CC polypeptide expression, particularly for regulating gene expression in
CC bacteria and for identifying proliferation-required genes or molecules
CC with potential antibiotic activity. The modified promoters are also
CC useful for replacing endogenous promoters to create cells with specific
CC regulatable genes. The present sequence is a fusion promoter sequence of
CC the invention. (Updated on 07-AUG-2003 to correct OS field.)
XX

SQ Sequence 80 BP; 29 A; 8 C; 8 G; 35 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
|||||
1 TCATAAAAAATTATTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 7
ACD13861
ID ACD13861 standard; DNA; 80 BP.

XX
AC ACD13861;

XX
DT 15-AUG-2003 (first entry)

DE Xyl-T5 fusion promoter #1.

XX
KM Promoter; ds; gram positive bacteria; *Staphylococcus aureus*;
KM *Enterococcus faecalis*; operator; xylO; teto; trpO; malO; lambda-clO;
KW cellular proliferation.

XX
OS Bacteriophage T5.
OS Unidentified.

XX
PN US2003027286-A1.

XX
PD 06-FEB-2003.

XX
PF 21-DEC-2001; 2001US-00032393.

XX
PR 06-SEP-2000; 2000US-0230335P.
PR 27-DEC-2000; 2000US-0259434P.

XX
PA (HASE/) HASELBECK R.

PA (WALL/) WALL D.
(GROS/) GROSS M.

XX
PI Haselbeck R, Wall D, Gross M;

XX
DR WPI; 2003-479541/45.

XX
PT New isolated nucleic acid comprising a fusion promoter having at least
PT one promoter that is modified to have altered activity in at least one
PT gram-positive organism, useful for regulating gene expression in
PT bacteria.

XX
PS Claim 24; Page 78; 142pp; English.

XX
CC The invention relates to an isolated nucleic acid comprising a fusion
CC promoter having at least one promoter that is modified to have altered
CC activity in at least one gram-positive organism (e.g. *Staphylococcus*
CC aureus or *Enterococcus faecalis*). The promoter is linked to at least one
CC operator selected from xylO, teto, trpO, malO and lambda-clO, which are
CC positioned such that the binding of at least one repressor to the
CC operator represses transcription from the fusion promoter. Also included
CC are a vector comprising the isolated nucleic acid, a host cell comprising
CC the nucleic acid. The fusion promoter is useful for identifying genes
CC involved in cellular proliferation, identifying a compound that reduces
CC the activity or level of a gene product required for proliferation of a
CC cell, inhibiting the activity or expression of a gene in an operon
CC required for proliferation, manufacturing an antibiotic, identifying a
CC gene that is required for proliferation of a prokaryotic cell,

CC identifying a compound that inhibits the proliferation of a prokaryotic
CC cell and regulating gene expression in bacteria. The present sequence is
CC a fusion promoter of the invention

XX Sequence 80 BP; 29 A; 8 C; 8 G; 35 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 8; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
Db 1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 8
ABK98589
ID ABK98589 standard; DNA; 94 BP.

XX ABK98589;

DT 07-AUG-2003 (revised)
DT 21-OCT-2002 (first entry)

DE Xy10-T5 fusion promoter oligonucleotide #1.

XX ss; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;
KW P59; P1P2; PL; xy10; tetO; trpO; malO; lambdaclO; cellular proliferation;
KW antibiotic.

XX Eubacteria.
OS Bacteriophage T5.
OS Synthetic.

XX WO200251982-A2.

XX 04-JUL-2002.

XX 21-DEC-2001; 2001WO-US050250.

XX 27-DEC-2000; 2000US-0259434P.

PR 06-SEP-2001; 2001US-00948993.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Wall D, Gross M;

DR WPI; 2002-575374/61.

PT Isolated nucleic acid comprises bacterial promoters modified to have
PT altered activity in at least one gram-positive organism, e.g. Bacillus
PT anthracis or Clostridium botulinum, useful for regulating gene expression
PT in bacteria.

PS Example 1; Page 81; 246pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion
CC promoter comprising at least one promoter that is modified to have
CC altered activity in at least one gram-positive organism, or comprising
CC T5, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting
CC of xy10, tetO, trpO, malO or lambdaclO, where at least one operator is
CC positioned so binding of a repressor to an operator represses
CC transcription from the fusion promoter. Also included are vectors and
CC host cells comprising the fusion promoters, a method of identifying genes
CC involved in cellular proliferation or required for proliferation of a
CC prokaryotic cell using the vector, a method of identifying compounds that
CC inhibit the proliferation of a prokaryotic cell using the vector, a
CC method of identifying a compound that reduces the activity or level of a
CC gene product required for proliferation of a cell using the vector, a
CC compound identified by the methods, a method of inhibiting the activity
CC or expression of a gene in an operon required for proliferation using the
CC vector, manufacturing an antibiotic comprising using the vector or cell
CC and identifying a nucleic acid with promoter activity in Enterococcus

CC faecalis. The fusion promoters are useful for regulating nucleic acid or
CC polypeptide expression, particularly for regulating gene expression in
CC bacteria and for identifying proliferation-regulated genes or molecules
CC with potential antibiotic activity. The modified promoters are also
CC useful for replacing endogenous promoters to create cells with specific
CC regulatable genes. The present sequence is an oligonucleotide used to
CC construct a fusion promoter sequence of the invention. (Updated on 07-AUG
CC -2003 to correct OS field.)

XX Sequence 94 BP; 35 A; 10 C; 12 G; 37 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
Db 6 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 56

RESULT 9
ACD13840
ID ACD13840 standard; DNA; 94 BP.

XX ACD13840;

DT 15-AUG-2003 (first entry)

DE Oligonucleotide Xyl-T5.

XX Promoter; ss; gram positive bacteria; Staphylococcus aureus;
KW Enterococcus faecalis; operator; xy10; tetO; trpO; malO; lambda-clO;
KW cellular proliferation.

XX Escherichia coli.
OS Synthetic.
OS Unidentified.

XX US2003027286-A1.

XX 06-FEB-2003.

XX 21-DEC-2001; 2001US-00032393.

XX 06-SEP-2000; 2000US-0230335P.

PR 27-DEC-2000; 2000US-0259434P.

PA (HASE/) HASELBECK R.

PA (WALL/) WALL D.

PI Haselbeck R, Wall D, Gross M;

DR WPI; 2003-479541/45.

PT New isolated nucleic acid comprising a fusion promoter having at least
PT one promoter that is modified to have altered activity in at least one
PT gram-positive organism, useful for regulating gene expression in
PT bacteria.

PS Example 1; Page 28; 142pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion
CC promoter having at least one promoter that is modified to have altered
CC activity in at least one gram-positive organism (e.g. Staphylococcus
CC aureus or Enterococcus faecalis). The promoter is linked to at least one
CC operator selected from xy10, tetO, trpO, malO and lambda-clO, which are
CC positioned such that the binding of at least one repressor to the
CC operator represses transcription from the fusion promoter. Also included
CC are a vector comprising the isolated nucleic acid, a host cell comprising
CC the nucleic acid. The fusion promoter is useful for identifying genes
CC involved in cellular proliferation, identifying a compound that reduces
CC the activity or level of a gene product required for proliferation of a

CC cell, inhibiting the activity or expression of a gene in an operon
CC required for proliferation, manufacturing an antibiotic, identifying a
CC gene that is required for proliferation of a prokaryotic cell,
CC identifying a compound that inhibits the proliferation of a prokaryotic
CC cell and regulating gene expression in bacteria. The present sequence is
CC an oligonucleotide used in the construction of a fusion promoter of the
CC invention
CC
SQ Sequence 94 BP; 35 A; 10 C; 12 G; 37 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 8; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATAAAAAATTATTTGCTTCAGGAAAAATTTTCTGTATAATAGATTCA 51
|||||
Db 6 TCATAAAAAATTATTTGCTTCAGGAAAAATTTTCTGTATAATAGATTCA 56

RESULT 10
ABK98590/c
ID ABK98590 standard; DNA; 95 BP.

XX ABK98590;
AC
XX 07-AUG-2003 (revised)
DT 21-OCT-2002 (first entry)
XX
DE XylO-T5 fusion promoter oligonucleotide #2.

XX ss; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;
KW P59; Plp2; Pl; xylO; teto; trpO; malO; lambdaclO; cellular proliferation;
KW antibiotic.
XX
OS Eubacteria.
OS Bacteriophage T5.
OS Synthetic.

XX WO200251982-A2.
PN
XX 04-JUL-2002.
PD
XX 21-DEC-2001; 2001WO-US050250.
PF
XX 27-DEC-2000; 2000US-0259434P.
PR 06-SEP-2001; 2001US-00948993.

XX (ELIT-) ELITRA PHARM INC.
PA
XX Haselbeck R, Wall D, Gross M;
PI
XX WPI; 2002-575374/61.
DR
XX
XX

PT Isolated nucleic acid comprises bacterial promoters modified to have
PT altered activity in at least one gram-positive organism, e.g. Bacillus
PT anthracis or Clostridium botulinum, useful for regulating gene expression
PT in bacteria.
PT
XX

PS Example 1; Page 81; 246pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion
CC promoter comprising at least one promoter that is modified to have
CC altered activity in at least one gram-positive organism, or comprising
CC T5, CP25, P32, P59, Plp2 or Pl linked to at least one operator consisting
CC of xylO, teto, trpO, malO or lambdaclO, where at least one operator is
CC positioned so binding of a repressor to an operator represses
CC transcription from the fusion promoter. Also included are vectors and
CC host cells comprising the fusion promoters, a method of identifying genes
CC involved in cellular proliferation or required for proliferation of a
CC prokaryotic cell using the vector, a method of identifying compounds that
CC inhibit the proliferation of a prokaryotic cell using the vector, a
CC method of identifying a compound that reduces the activity or level of a
CC gene product required for proliferation of a cell using the vector, a

CC compound identified by the methods, a method of inhibiting the activity
CC or expression of a gene in an operon required for proliferation using the
CC vector, manufacturing an antibiotic comprising using the vector or cell
CC and identifying a nucleic acid with promoter activity in Enterococcus
CC faecalis. The fusion promoters are useful for regulating nucleic acid or
CC polypeptide expression, particularly for regulating gene expression in
CC bacteria and for identifying proliferation-regulated genes or molecules
CC with potential antibiotic activity. The modified promoters are also
CC useful for replacing endogenous promoters to create cells with specific
CC regulatable genes. The present sequence is an oligonucleotide used to
CC construct a fusion promoter sequence of the invention. (Updated on 07-AUG
CC -2003 to correct OS field.)
CC
XX
SQ Sequence 95 BP; 39 A; 11 C; 9 G; 36 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATAAAAAATTATTTGCTTCAGGAAAAATTTTCTGTATAATAGATTCA 51
|||||
Db 93 TCATAAAAAATTATTTGCTTCAGGAAAAATTTTCTGTATAATAGATTCA 43

RESULT 11
ACD13841/c
ID ACD13841 standard; DNA; 95 BP.

XX ACD13841;
AC
XX 15-AUG-2003 (first entry)
DT
XX Oligonucleotide Xyl-T5 complement.
DE
XX

KW Promoter; ss; gram positive bacteria; Staphylococcus aureus;
KW Enterococcus faecalis; operator; xylO; teto; trpO; malO; lambda-clO;
KW cellular proliferation.
XX

XX Escherichia coli.
OS Synthetic.
OS Unidentified.

XX US2003027286-A1.

XX 06-FEB-2003.

XX 21-DEC-2001; 2001US-00032393.

XX 06-SEP-2000; 2000US-0230335P.
PR 27-DEC-2000; 2000US-0259434P.

XX (HASE/) HASELBECK R.
PA (WALL/) WALL D.
PA (GROS/) GROSS M.

PI Haselbeck R, Wall D, Gross M;
PI
XX WPI; 2003-479541/45.
DR
XX

PT New isolated nucleic acid comprising a fusion promoter having at least
PT one promoter that is modified to have altered activity in at least one
PT gram-positive organism, useful for regulating gene expression in
PT bacteria.
PT
XX

PS Example 1; Page 28; 142pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion
CC promoter having at least one promoter that is modified to have altered
CC activity in at least one gram-positive organism (e.g. Staphylococcus
CC aureus or Enterococcus faecalis). The promoter is linked to at least one
CC operator selected from xylO, teto, trpO, malO and lambda-clO, which are
CC positioned such that the binding of at least one repressor to the
CC operator represses transcription from the fusion promoter. Also included

CC are a vector comprising the isolated nucleic acid, a host cell comprising
CC the nucleic acid. The fusion promoter is useful for identifying genes
CC involved in cellular proliferation, identifying a compound that reduces
CC the activity or level of a gene product required for proliferation of a
CC cell, inhibiting the activity or expression of a gene in an operon
CC required for proliferation, manufacturing an antibiotic, identifying a
CC gene that is required for proliferation of a prokaryotic cell,
CC identifying a compound that inhibits the proliferation of a prokaryotic
CC cell and regulating gene expression in bacteria. The present sequence is
CC an oligonucleotide used in the construction of a fusion promoter of the
CC invention

SQ Sequence 95 BP; 39 A; 11 C; 9 G; 36 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 8; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 93 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 43

RESULT 12

ID AAN91066 standard; DNA; 200 BP.

AC AAN91066;

DT 03-OCT-2002 (revised)

DT 09-JUL-1989 (first entry)

DE Control sequence N25*/O.

KM DNA expression control; ss.

OS Synthetic.

PN EP303925-A.

PD 22-FEB-1989.

PF 08-AUG-1988; 88EP-00112864.

PR 17-AUG-1987; 87CH-00003152.

PA (HOFF) HOFFMANN-LA ROCHE AG.

PI Bujard H, Lanzer M;

DR WPI; 1989-055375/08.

PT Deoxyribonucleic acid data expression control sequences - comprising
PT promoter and operator-repressor-sequences.

PS Disclosure; Fig 7; 65pp; German.

CC This sequence is the operator N25*/O and is inserted into plasmid pDS3
CC for expression of a variety of proteins from pro- and eukaryotic sources.
CC It is prepared as a DNA XhoI/EcoRI frag. This operator/repressor sequence
CC has a high complexing rate and gives good repressability. See also
CC N910601-3,5,7,8 and AAN91070. (Updated on 03-OCT-2002 to add missing OS
CC field.)

SQ Sequence 200 BP; 64 A; 33 C; 37 G; 66 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 124 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 174

RESULT 13

ID AAN60262 standard; DNA; 202 BP.

AC AAN60262;

DT 29-MAY-1991 (first entry)

DE Plasmid pDS2/P promoter/operator fusion (PN25x/O).

KM DHFR; chloramphenicol acetyl-transferase; malaria SA; IL-2; IFN; insulin;

KW tPA; renin; ds.

OS Synthetic.

PN EP186069-A.

PD 02-JUL-1986.

PF 13-DEC-1985; 85EP-00115921.

PR 17-DEC-1984; 84GB-00031818.

PA (HOFF) HOFFMANN-LA ROCHE AG.

PI Bujard H, Scuber D;

DR WPI; 1986-170629/27.

PT Expression control DNA sequence - comprising T5 promoter combined with
PT DNA sequence which permits control of promoter activity.

PS Disclosure; Fig 7; 26pp; English.

CC Plasmid vectors of the pDS1 family may be used to express a sequence
CC under the control of the colliphage T5 promoter, and one or more
CC sequences which allow control of the promoter. Sequences expressed
CC include products such as dihydrofolate reductase; chloramphenicol acetyl-
CC transferase; malaria surface antigen; IL-2; IFN-alpha, -beta and -gamma;
CC insulin; growth hormones; tPA; human renin etc

SQ Sequence 202 BP; 64 A; 34 C; 37 G; 67 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 124 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 174

RESULT 14

ID AAN91067 standard; DNA; 207 BP.

AC AAN91067;

DT 03-OCT-2002 (revised)

DT 09-JUL-1989 (first entry)

DE Control sequence N25OP29.

KM DNA expression control; ss.

OS Synthetic.

PN EP303925-A.

PD 22-FEB-1989.

XX

PF 08-AUG-1988; 88EP-00112864.
XX 17-AUG-1987; 87CH-00003152.
PR (HOFF) HOFFMANN-LA ROCHE AG.
XX
XX Bujard H, Lanzner M;
XX WPI; 1989-055375/08.
DR
XX Deoxyribonucleic acid data expression control sequences - comprising
PT Promoter and operator-repressor-sequences.
XX
XX Disclosure; Fig 7; 65pp; German.
PS
XX Sequence is the operator/promoter N25OP29 and is inserted into plasmid
CC PDS3 for expression of a variety of proteins from pro- and eukaryotic
CC sources. It is prepared as a DNA XhoI/EcoRI frag. It comprises a low
CC signal strength/high promoter strength promoter and high complexing rate
CC operator/repressor. It gives high transcription and translation efficienc
CC ies and good repressability. See also AAN91060, 1,2,3,5,6,8 and AAN91070.
CC (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ Sequence 207 BP; 66 A; 37 C; 37 G; 67 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
124 TCATAAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 174
DB

RESULT 15
AAN60259
ID AAN60259 standard; DNA; 247 BP.
XX
AC AAN60259;
XX
DT 27-AUG-2003 (revised)
DT 29-MAY-1991 (first entry)
XX
DE Coliphage PN25 Pre-early promoter of plasmid pDS1, PN25, tol+.
XX
XX DHFR; chloramphenicol acetyl-transferase; malaria SA; IL-2; IFN; insulin;
KM tPA; renin.
XX
OS Unidentified.
XX
XX EP186069-A.
XX
PD 02-JUL-1986.
XX
XX 13-DEC-1985; 85EP-00115921.
PF
XX 17-DEC-1984; 84GB-00031818.
PR
XX (HOFF) HOFFMANN-LA ROCHE AG.
PA
XX Bujard H, Stuber D;
PI
XX WPI; 1986-170629/27.
DR
XX
XX Expression control DNA sequence - comprising T5 promoter combined with
PT DNA sequence which permits control of promoter activity.
XX
XX Disclosure; Fig 2; 26pp; English.
PS
XX Plasmid vectors of the pDS1 family may be used to express a sequence
CC under the control of a coliphage T5 promoter eg. the coliphage PN25 pre-
CC early promoter, and one or more sequences which allow control of the
CC promoter. Sequences expressed include products such as dihydrofolate

CC reductase; chloramphenicol acetyl-transferase; malaria surface antigen;
CC IL-2; IFN-alpha, -beta and -gamma; insulin; growth hormones; tPA; human
CC renin etc. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 247 BP; 79 A; 44 C; 48 G; 76 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
124 TCATAAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 174
DB

Search completed: May 4, 2004, 10:37:46
Job time : 391 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 09:55:29 ; Search time 76 Seconds
(without alignments)
372.401 Million cell updates/sec

Title: US-10-032-393-36
Perfect score: 51
Sequence: 1 tcataaaaatttattgtct.....ttttctgtataatagattca 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33.4	65.5	125	1	US-08-400-864-2 Sequence 2, Appli
2	33	64.7	171	2	US-08-892-272-3 Sequence 3, Appli
3	33	64.7	757	2	US-08-892-272-1 Sequence 1, Appli
4	31.4	61.6	49	1	US-08-400-864-1 Sequence 1, Appli
5	28.6	56.1	247	3	US-09-344-888A-24 Sequence 24, Appli
6	28.6	56.1	250	3	US-08-840-466A-25 Sequence 25, Appli
7	28.6	56.1	250	4	US-09-696-188B-25 Sequence 25, Appli
8	28.6	56.1	520	1	US-08-268-348A-7 Sequence 7, Appli
9	28.6	56.1	520	1	US-08-268-348A-9 Sequence 9, Appli
10	28.6	56.1	1106	1	US-08-041-648-4 Sequence 4, Appli
11	28.6	56.1	1106	1	US-08-041-648-6 Sequence 6, Appli
12	28.6	56.1	1106	1	US-08-041-648-8 Sequence 8, Appli
13	28.6	56.1	3462	4	US-09-742-373-10 Sequence 10, Appli
14	28.6	56.1	3977	1	US-07-794-400-2 Sequence 2, Appli
15	28.6	56.1	3977	1	US-07-794-400-13 Sequence 13, Appli
16	28.6	56.1	3977	1	US-08-041-648-1 Sequence 1, Appli
17	28.6	56.1	3977	1	US-08-217-529-1 Sequence 1, Appli
18	28.6	56.1	3977	1	US-08-397-470-2 Sequence 2, Appli
19	28.6	56.1	3977	1	US-08-397-470-13 Sequence 13, Appli
20	28.6	56.1	4202	4	US-09-483-419-2 Sequence 2, Appli
21	28.6	56.1	4491	4	US-09-837-863-23 Sequence 23, Appli
22	28.6	56.1	4755	4	US-09-837-863-24 Sequence 24, Appli
23	28.6	56.1	6501	4	US-09-767-515-1 Sequence 1, Appli
24	28.6	56.1	6501	4	US-09-767-515-2 Sequence 2, Appli
25	27.2	53.3	1830121	4	US-09-557-884-1 Sequence 1, Appli
26	27.2	53.3	1830121	4	US-09-643-990A-1 Sequence 1, Appli
27	27	52.9	4519	4	US-08-976-259-6 Sequence 6, Appli

C	28	25.4	49.8	843	4	US-09-134-001C-2669	Sequence 2669, Ap
	29	25	49.0	92407	4	US-09-596-002-36	Sequence 36, Appli
	30	25	49.0	640681	4	US-09-790-988-1	Sequence 1, Appli
C	31	24.8	48.6	6152	3	US-08-973-462-1	Sequence 1, Appli
C	32	24.4	47.8	2728	4	US-09-620-312D-572	Sequence 572, App
C	33	24.4	47.8	15788	4	US-09-920-759-13	Sequence 13, Appli
	34	24.4	47.8	193303	4	US-09-497-855A-37	Sequence 37, Appli
	35	24.4	47.8	193303	4	US-09-497-855A-44	Sequence 44, Appli
	36	24.2	47.5	1650	4	US-09-905-125A-254	Sequence 254, App
	37	24.2	47.5	1650	4	US-09-905-125A-254	Sequence 254, App
	38	24.2	47.5	1650	4	US-09-902-775A-254	Sequence 254, App
C	39	24	47.1	640681	4	US-09-790-988-1	Sequence 1, Appli
	40	23.8	46.7	786431	4	US-09-751-389-3	Sequence 3, Appli
	41	23.6	46.3	2243	1	US-07-995-657-1	Sequence 1, Appli
	42	23.6	46.3	2243	1	US-08-474-587-1	Sequence 1, Appli
	43	23.6	46.3	2349	2	US-08-974-546-2	Sequence 2, Appli
	44	23.6	46.3	6113	4	US-10-204-708-14	Sequence 14, Appli
C	45	23.6	46.3	112132	4	US-09-741-150-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-400-864-2
; Sequence 2, Application US/08400864
; Patent No. 5721137
; GENERAL INFORMATION:
; APPLICANT: FRASCOTTI, GIANNI
; APPLICANT: GRANDI, GUIDO
; TITLE OF INVENTION: PLASMID VECTOR AND ITS USE FOR THE
; TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,864
; FILING DATE: 08-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT 000727 MI94A
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2264-083-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-400-864-2

Query Match 65.5%; Score 33.4; DB 1; Length 125;
Best Local Similarity 95.7%; Pred. No. 0.16;
Matches 45; Conservative 0; Mismatches 1; Indels 1; Gaps 1;


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT 000727 MI94A
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2264-083-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-400-864-1

Query Match      61.6%; Score 31.4; DB 1; Length 49;
Best Local Similarity 95.6%; Pred. No. 0.57;
Matches 43; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 6 AAAAATTATTTGCTTTCAGGAAA-TTTTCTGTATATAGATT 49
    |||||||
Db 5 AAAAATTATTTGCTTTCAGGAAA-TTTTCTGTATATAGATT 49

RESULT 5
US-09-344-888A-24
; Sequence 24, Application US/09344888A
; Patent No. 6291245
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard
; APPLICANT: Schantz, Christian
; TITLE OF INVENTION: New Host-Vector System
; FILE REFERENCE: CD20315
; CURRENT APPLICATION NUMBER: US/09/344,888A
; CURRENT FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: EP98113156.8
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EP98119078.8
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:expression cassette
; US-09-344-888A-24

Query Match      56.1%; Score 28.6; DB 3; Length 247;
Best Local Similarity 72.5%; Pred. No. 3.8;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTTCAGGAAAATTTTCTGTATATAGATTCA 51
    |||||||
Db 10 TCATATAAAATTATTTGCTTTCAGGATAACAATATATAGATTCA 60

RESULT 6
US-08-840-466A-25
; Sequence 25, Application US/08840466A
; Patent No. 6261561
; GENERAL INFORMATION:
; APPLICANT: Stewart, C. Neal
; APPLICANT: McKee, Marian L.
; APPLICANT: O'Brien, Alison D.
; APPLICANT: Wachtel, Marian R.
; TITLE OF INVENTION: Method Of Stimulating An Immune Response
```

```

; By Administration Of Host Organisms That Express Intimin
; Alone Or As A Fusion Protein With One Or More Other
; Antigens.
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,466A
; FILING DATE: 18-Apr-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Boone, Laurel S.
; REGISTRATION NUMBER: 43,505
; REFERENCE/DOCKET NUMBER: 04995.0029-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
; US-08-840-466A-25

Query Match      56.1%; Score 28.6; DB 3; Length 250;
Best Local Similarity 72.5%; Pred. No. 3.8;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTTCAGGAAAATTTTCTGTATATAGATTCA 51
    |||||||
Db 10 TCATATAAAATTATTTGCTTTCAGGATAACAATATATAGATTCA 60

RESULT 7
US-09-696-188B-25
; Sequence 25, Application US/09696188B
; Patent No. 6406885
; GENERAL INFORMATION:
; APPLICANT: Stewart, C. Neal
; APPLICANT: McKee, Marian L.
; APPLICANT: O'Brien, Alison D.
; APPLICANT: Wachtel, Marian R.
; TITLE OF INVENTION: Method Of Stimulating An Immune Response
; By Administration Of Host Organisms That Express Intimin
; Alone Or As A Fusion Protein With One Or More Other
; Antigens.
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```


SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/696,188B
FILING DATE: 26-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/840,466
FILING DATE: 1997-04-18
ATTORNEY/AGENT INFORMATION:
NAME: Boone, Laurel S.
REGISTRATION NUMBER: 43,505
REFERENCE/DOCKET NUMBER: 04995.0029-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-696-188B-25

Query Match 56.1%; Score 28.6; DB 4; Length 250;
Best Local Similarity 72.5%; Pred. No. 3.8;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
|||||
Db 10 TCATATAAAATTATTTGCTTCAGGATTAACAATTATATAGATTCA 60

RESULT 8

US-08-268-348A-7

Sequence 7, Application US/08268348A
Patent No. 5750374

GENERAL INFORMATION:

APPLICANT: Dobeli, Heinz

APPLICANT: Draeger, Nicholas

APPLICANT: Trotman, Gerda H

APPLICANT: Jakob, Peter

APPLICANT: Stuber, Dietrich

TITLE OF INVENTION: Process for Producing Hydrophobic

TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in

TITLE OF INVENTION: Producing Same

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann-La Roche Inc.

STREET: 340 Kingsland Street

CITY: Nutley

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/268,348A

FILING DATE: 29-JUN-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 93110755.1

FILING DATE: 06-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Parise, John P.

REGISTRATION NUMBER: 34,403

REFERENCE/DOCKET NUMBER: 4105/157

TELECOMMUNICATION INFORMATION:

TELEPHONE: (201) 235-6326

TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 115..516
OTHER INFORMATION: /product= "Amyloid Protein AA"
US-08-268-348A-7

Query Match 56.1%; Score 28.6; DB 1; Length 520;
Best Local Similarity 72.5%; Pred. No. 4;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
|||||
Db 10 TCATATAAAATTATTTGCTTCAGGATTAACAATTATATAGATTCA 60

RESULT 9

US-08-268-348A-9

Sequence 9, Application US/08268348A
Patent No. 5750374

GENERAL INFORMATION:

APPLICANT: Dobeli, Heinz

APPLICANT: Draeger, Nicholas

APPLICANT: Trotman, Gerda H

APPLICANT: Jakob, Peter

APPLICANT: Stuber, Dietrich

TITLE OF INVENTION: Process for Producing Hydrophobic

TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in

TITLE OF INVENTION: Producing Same

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann-La Roche Inc.

STREET: 340 Kingsland Street

CITY: Nutley

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/268,348A

FILING DATE: 29-JUN-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 93110755.1

FILING DATE: 06-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Parise, John P.

REGISTRATION NUMBER: 34,403

REFERENCE/DOCKET NUMBER: 4105/157

TELECOMMUNICATION INFORMATION:

TELEPHONE: (201) 235-6326

TELEFAX: (201) 235-3500

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 520 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 115..516

OTHER INFORMATION: /product= "Amyloid Protein AA"
US-08-268-346A-9

Query Match 56.1%; Score 28.6; DB 1; Length 520;
Best Local Similarity 72.5%; Pred. No. 4;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
10 TCATAAAAAATTATTTGCTTTCAGGCGGATACATTATATAGATTCA 60

RESULT 10
US-08-041-648-4

; Sequence 4, Application US/08041648

; Patent No. 5486463

; GENERAL INFORMATION:

; APPLICANT: Lesslauer, Werner

; APPLICANT: L tscher, Hansruedi

; APPLICANT: St ber, Dietrich

; TITLE OF INVENTION: TNF-MUTAINS

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07110-1199

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/041,648

; FILING DATE: 1-APR-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 92810249.0

; FILING DATE: 2-APR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Roseman, Catherine R.

; REGISTRATION NUMBER: 34240

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-6208

; TELEFAX: (201) 235-3500

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1106 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: Plasmid pDS56/RBSII, SphI-TNFalpha (Trp32)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 994..1104

US-08-041-648-4

Query Match 56.1%; Score 28.6; DB 1; Length 1106;
Best Local Similarity 72.5%; Pred. No. 4.1;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
886 TCATAAAAAATTATTTGCTTTCAGGCGGATACATTATATAGATTCA 936

RESULT 11
US-08-041-648-6

; Sequence 6, Application US/08041648

; Patent No. 5486463

; GENERAL INFORMATION:

; APPLICANT: Lesslauer, Werner

; APPLICANT: L tscher, Hansruedi

; APPLICANT: St ber, Dietrich

; TITLE OF INVENTION: TNF-MUTAINS

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07110-1199

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/041,648

; FILING DATE: 1-APR-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 92810249.0

; FILING DATE: 2-APR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Roseman, Catherine R.

; REGISTRATION NUMBER: 34240

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-6208

; TELEFAX: (201) 235-3500

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1106 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; ORGANISM: Plasmid pDS56/RBSII, SphI-TNFalpha (Ser29)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 994..1104

US-08-041-648-6

Query Match 56.1%; Score 28.6; DB 1; Length 1106;
Best Local Similarity 72.5%; Pred. No. 4.1;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTTGCTTTCAGGAAATTTTCTGTATATAGATTCA 51
886 TCATAAAAAATTATTTGCTTTCAGGCGGATACATTATATAGATTCA 936

RESULT 12

US-08-041-648-8

; Sequence 8, Application US/08041648

; Patent No. 5486463

; GENERAL INFORMATION:

; APPLICANT: Lesslauer, Werner

; APPLICANT: L tscher, Hansruedi

; APPLICANT: St ber, Dietrich

; TITLE OF INVENTION: TNF-MUTAINS

; NUMBER OF SEQUENCES: 17

ପରସ୍ପର

```

TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Vectorsequence ,
OTHER INFORMATION: DNA
US-09-742-373-10

Query Match          56.1%; Score 28.6; DB 4; Length 3462;
Best Local Similarity 72.5%; Pred. No. 4.3;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

CY 1 TCATAAAAAATTATTTGCTTTCAGAAATTTTCTGTATATAGATTCA 51
    |||||||
10 TCATAAAAAATTATTTGCTTGTGAGCGGATACAAATTATAATAGATTCA 60

RESULT 14
US-07-794-400-2
; Sequence 2, Application US/07794400
; Patent No. 5422104
; GENERAL INFORMATION:
; APPLICANT: Fiers, W.
; APPLICANT: Tavernier, J.
; APPLICANT: Van Ostade, X.
; TITLE OF INVENTION: TNF-Mutetins
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/794,400
; FILING DATE: 19911120
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90810901.0
; FILING DATE: 21-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Krovatin, William
; REGISTRATION NUMBER: 33256
; REFERENCE/DOCKET NUMBER: 4105/136-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4387
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3977 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant plasmid)
; IMMEDIATE SOURCE:
; CLONE: pDS56/RBSII,SphI-TNF-alpha
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..591
; US-07-794-400-2

Query Match          56.1%; Score 28.6; DB 1; Length 3977;
Best Local Similarity 72.5%; Pred. No. 4.3;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

CY 1 TCATAAAAAATTATTTGCTTTCAGAAATTTTCTGTATATAGATTCA 51
    |||||||
10 TCATAAAAAATTATTTGCTTGTGAGCGGATACAAATTATAATAGATTCA 60

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 11:34:18 ; Search time 403 Seconds
(without alignments)
572.448 Million cell updates/sec

Title: US-10-032-393-36
Perfect score: 51
Sequence: 1 tcaataaaatttatttgcgtct.....ttttctgtataatagattca 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	51	100.0	51	15	US-10-032-393-36 Sequence 36, Appl
2	51	100.0	72	9	US-09-788-297-6 Sequence 6, Appl
3	51	100.0	80	9	US-09-788-297-7 Sequence 7, Appl
4	51	100.0	80	15	US-10-032-393-26 Sequence 26, Appl
5	51	100.0	94	15	US-10-032-393-5 Sequence 5, Appl
6	51	100.0	95	15	US-10-032-393-6 Sequence 6, Appl
7	51	100.0	556	15	US-10-032-393-1 Sequence 1, Appl
8	51	100.0	5302	16	US-10-032-393-1 Sequence 2, Appl
9	51	100.0	5767	16	US-10-385-415-4 Sequence 4, Appl
10	51	100.0	6852	15	US-10-032-393-16 Sequence 16, Appl
11	50	98.0	1246	9	US-09-815-242-3094 Sequence 3094, Ap
12	43	84.3	51	15	US-10-032-393-37 Sequence 37, Appl
13	43	84.3	80	15	US-10-032-393-27 Sequence 27, Appl
14	43	84.3	94	15	US-10-032-393-22 Sequence 22, Appl

c	15	43	84.3	95	15	US-10-032-393-23	Sequence 23, Appl
	16	43	84.3	6852	15	US-10-032-393-53	Sequence 53, Appl
	17	29.8	58.4	113515	15	US-10-311-455-2148	Sequence 2148, Ap
	18	28.6	56.1	82	15	US-10-288-858-13	Sequence 13, Appl
	19	28.6	56.1	94	15	US-10-032-393-12	Sequence 12, Appl
c	20	28.6	56.1	95	15	US-10-032-393-13	Sequence 13, Appl
	21	28.6	56.1	134	15	US-10-284-083-3	Sequence 3, Appl
	22	28.6	56.1	138	15	US-10-284-083-4	Sequence 4, Appl
	23	28.6	56.1	138	15	US-10-284-083-5	Sequence 5, Appl
	24	28.6	56.1	162	9	US-09-828-523A-97	Sequence 97, Appl
	25	28.6	56.1	162	10	US-09-966-521-89	Sequence 89, Appl
	26	28.6	56.1	162	15	US-10-429-094-89	Sequence 89, Appl
	27	28.6	56.1	166	15	US-10-284-083-2	Sequence 2, Appl
	28	28.6	56.1	170	15	US-10-284-083-6	Sequence 6, Appl
	29	28.6	56.1	250	8	US-08-837-459-25	Sequence 25, Appl
	30	28.6	56.1	250	15	US-10-150-058-25	Sequence 25, Appl
	31	28.6	56.1	2220	15	US-10-288-858-5	Sequence 5, Appl
	32	28.6	56.1	3420	16	US-10-385-415-1	Sequence 1, Appl
	33	28.6	56.1	3462	9	US-09-742-373-10	Sequence 10, Appl
	34	28.6	56.1	3462	15	US-10-394-058-10	Sequence 10, Appl
	35	28.6	56.1	3876	16	US-10-385-415-27	Sequence 27, Appl
	36	28.6	56.1	3876	16	US-10-385-415-28	Sequence 28, Appl
	37	28.6	56.1	3879	16	US-10-385-415-5	Sequence 5, Appl
	38	28.6	56.1	3879	16	US-10-385-415-6	Sequence 6, Appl
	39	28.6	56.1	3879	16	US-10-385-415-7	Sequence 7, Appl
	40	28.6	56.1	3879	16	US-10-385-415-21	Sequence 21, Appl
	41	28.6	56.1	3879	16	US-10-385-415-25	Sequence 25, Appl
	42	28.6	56.1	3879	16	US-10-385-415-26	Sequence 26, Appl
	43	28.6	56.1	3885	16	US-10-385-415-3	Sequence 3, Appl
	44	28.6	56.1	3897	16	US-10-385-415-20	Sequence 20, Appl
	45	28.6	56.1	3900	16	US-10-385-415-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-10-032-393-36
; Sequence 36, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: EITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyl-T5 promoter sequence
US-10-032-393-36

Query Match 100.0%; Score 51; DB 15; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTTCTGTATTAATAGATTCA 51
DB 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTTCTGTATTAATAGATTCA 51

RESULT 2

```
US-09-788-297-6
; Sequence 6, Application US/09788297
; Patent No. US20020094516A1
; GENERAL INFORMATION:
; APPLICANT: Calos, Michele P.
; APPLICANT: Scilimenti, Christopher R.
; TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION
; FILE REFERENCE: 8400-0011
; CURRENT APPLICATION NUMBER: US/09/788,297
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 72
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 1
US-09-788-297-6
```

```
Query Match          100.0%; Score 51; DB 9; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
          |||
Db       2 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 52
```

RESULT 3

```
US-09-788-297-7/c
; Sequence 7, Application US/09788297
; Patent No. US20020094516A1
; GENERAL INFORMATION:
; APPLICANT: Calos, Michele P.
; APPLICANT: Scilimenti, Christopher R.
; TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION
; FILE REFERENCE: 8400-0011
; CURRENT APPLICATION NUMBER: US/09/788,297
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 2
US-09-788-297-7
```

```
Query Match          100.0%; Score 51; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
          |||
Db       75 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 25
```

RESULT 4

```
US-10-032-393-26
; Sequence 26, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
```

```
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyl-T5 fusion sequence
US-10-032-393-26
```

```
Query Match          100.0%; Score 51; DB 15; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
          |||
Db       1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
```

RESULT 5

```
US-10-032-393-5
; Sequence 5, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 94
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Xyl-T5
US-10-032-393-5
```

```
Query Match          100.0%; Score 51; DB 15; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
          |||
Db       6 TCATATAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 56
```

RESULT 6

```
US-10-032-393-6/c
; Sequence 6, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
```



```
APPLICANT: Gross, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010A
CURRENT APPLICATION NUMBER: US/10/032,393
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: 09/948,993
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 6852
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Vector pXyl-T5 p15a
US-10-032-393-16
```

```
Query Match      100.0%; Score 51; DB 15; Length 6852;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
Db      1367 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 1417
```

```
RESULT 11
US-09-815-242-3094
; Sequence 3094, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3094
; LENGTH: 1246
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1246)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-242-3094
```

```
Query Match      98.0%; Score 50; DB 9; Length 1246;
Best Local Similarity 98.0%; Pred. No. 6e-05;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
Db      13 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 63
```

```
RESULT 12
US-10-032-393-37
; Sequence 37, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyl-T5-DD promoter sequence
US-10-032-393-37
```

```
Query Match      84.3%; Score 43; DB 15; Length 51;
Best Local Similarity 90.2%; Pred. No. 0.003;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
Db      1 TCATAAAAAATTATTACATCAGGAAATTTTCTGTATATAGATTCA 51
```

```
RESULT 13
US-10-032-393-27
; Sequence 27, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyl-T5-DD fusion sequence
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 10:31:40 ; Search time 3247 Seconds
(without alignments)
552.662 Million cell updates/sec

Title: US-10-032-393-36
Perfect score: 51
Sequence: 1 tcataaaaaatttatttgc.....tttctgtataatagattca 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*

2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*

3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*

4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*

5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*

6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*

7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*

8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*

9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*

10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*

11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*

12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*

13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*

14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*

15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*

16: /cgn2_6/ptodata/2/pna/US092A_COMB.seq:*

17: /cgn2_6/ptodata/2/pna/US092B_COMB.seq:*

18: /cgn2_6/ptodata/2/pna/US093A_COMB.seq:*

19: /cgn2_6/ptodata/2/pna/US093B_COMB.seq:*

20: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*

21: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*

22: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*

23: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*

24: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*

25: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*

26: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*

27: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*

28: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*

29: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*

30: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*

31: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*

32: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*

33: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*

34: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*

35: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*

36: /cgn2_6/ptodata/2/pna/US098D_COMB.seq:*

37: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*

38: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*

39: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*

40: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*

41: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*

42: /cgn2_6/ptodata/2/pna/US099F_COMB.seq:*

43: /cgn2_6/ptodata/2/pna/US100A_COMB.seq:*

44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*

45: /cgn2_6/ptodata/2/pna/US101A_COMB.seq:*

46: /cgn2_6/ptodata/2/pna/US101B_COMB.seq:*

47: /cgn2_6/ptodata/2/pna/US102A_COMB.seq:*

48: /cgn2_6/ptodata/2/pna/US102B_COMB.seq:*

49: /cgn2_6/ptodata/2/pna/US103A_COMB.seq:*

50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq:*

51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq:*

52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq:*

53: /cgn2_6/ptodata/2/pna/US106_COMB.seq:*

54: /cgn2_6/ptodata/2/pna/US107A_COMB.seq:*

55: /cgn2_6/ptodata/2/pna/US107B_COMB.seq:*

56: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*

57: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*

58: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*

59: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*

60: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*

61: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*

62: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*

63: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*

64: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*

65: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*

66: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*

67: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*

68: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*

69: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*

70: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*

71: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*

72: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*

73: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*

74: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*

75: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*

76: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*

77: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*

78: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*

79: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:*

80: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:*

81: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*

82: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*

83: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*

84: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*

85: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*

86: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*

87: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*

88: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*

89: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*

90: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*

91: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*

92: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*

93: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*

94: /cgn2_6/ptodata/2/pna/US6037_COMB.seq:*

95: /cgn2_6/ptodata/2/pna/US6038_COMB.seq:*

96: /cgn2_6/ptodata/2/pna/US6039_COMB.seq:*

97: /cgn2_6/ptodata/2/pna/US6040_COMB.seq:*

98: /cgn2_6/ptodata/2/pna/US6041_COMB.seq:*

99: /cgn2_6/ptodata/2/pna/US6042_COMB.seq:*

100: /cgn2_6/ptodata/2/pna/US6043_COMB.seq:*

101: /cgn2_6/ptodata/2/pna/US6044_COMB.seq:*

102: /cgn2_6/ptodata/2/pna/US6045_COMB.seq:*

103: /cgn2_6/ptodata/2/pna/US6046_COMB.seq:*

104: /cgn2_6/ptodata/2/pna/US6047_COMB.seq:*

105: /cgn2_6/ptodata/2/pna/US6048_COMB.seq:*

106: /cgn2_6/ptodata/2/pna/US6049_COMB.seq:*

107: /cgn2_6/ptodata/2/pna/US6050_COMB.seq:*

108: /cgn2_6/ptodata/2/pna/US6051_COMB.seq:*

109: /cgn2_6/ptodata/2/pna/US6052_COMB.seq:*

110: /cgn2_6/ptodata/2/pna/US6053_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	100.0	51	1	PCT-US01-50250-36	Sequence 36, Appl
2	51	100.0	51	1	PCT-US01-50250A-36	Sequence 36, Appl
3	51	100.0	51	43	US-10-032-393-36	Sequence 36, Appl
4	51	100.0	72	1	PCT-US01-05269-6	Sequence 6, Appli
5	51	100.0	72	32	US-09-788-297-6	Sequence 6, Appli
6	51	100.0	80	1	PCT-US01-05269-7	Sequence 7, Appli
7	51	100.0	80	1	PCT-US01-50250-26	Sequence 26, Appl
8	51	100.0	80	1	PCT-US01-50250A-26	Sequence 26, Appl
9	51	100.0	80	32	US-09-788-297-7	Sequence 7, Appli
10	51	100.0	80	43	US-10-032-393-26	Sequence 26, Appl
11	51	100.0	94	1	PCT-US01-50250-5	Sequence 5, Appli
12	51	100.0	94	1	PCT-US01-50250A-5	Sequence 5, Appli
13	51	100.0	94	43	US-10-032-393-5	Sequence 5, Appli
14	51	100.0	94	82	US-60-259-434-3	Sequence 3, Appli
15	51	100.0	95	1	PCT-US01-50250-6	Sequence 6, Appli
16	51	100.0	95	1	PCT-US01-50250A-6	Sequence 6, Appli
17	51	100.0	95	43	US-10-032-393-6	Sequence 6, Appli
18	51	100.0	95	82	US-60-259-434-4	Sequence 99, Appl
19	51	100.0	265	76	US-60-206-848-99	Sequence 22, Appl
20	51	100.0	297	76	US-60-206-848-22	Sequence 1609, Ap
21	51	100.0	343	82	US-60-253-625-1609	Sequence 232, App
22	51	100.0	384	76	PCT-US01-50250-1	Sequence 1, Appli
23	51	100.0	556	1	PCT-US01-50250A-1	Sequence 1, Appli
24	51	100.0	556	43	US-10-032-393-1	Sequence 1, Appli
25	51	100.0	556	82	US-60-259-434-10	Sequence 10, Appl
26	51	100.0	5302	38	US-09-936-028-2	Sequence 2, Appli
27	51	100.0	5302	38	US-09-936-028A-2	Sequence 2, Appli
28	51	100.0	5302	50	US-10-385-415-2	Sequence 4, Appli
29	51	100.0	5302	38	US-09-936-028-4	Sequence 4, Appli
30	51	100.0	5767	38	US-09-936-028A-4	Sequence 4, Appli
31	51	100.0	5767	50	US-10-385-415-4	Sequence 16, Appl
32	51	100.0	6852	1	PCT-US01-50250-16	Sequence 16, Appl
33	51	100.0	6852	1	PCT-US01-50250A-16	Sequence 16, Appl
34	51	100.0	6852	43	US-10-032-393-16	Sequence 8, Appli
35	51	100.0	6858	82	US-60-259-434-8	Sequence 3094, Ap
36	50	98.0	1246	1	PCT-US02-03987-3094	Sequence 3094, Ap
37	50	98.0	1246	33	US-09-815-242-3094	Sequence 3094, Ap
38	50	98.0	1246	44	US-10-072-851-3094	Sequence 15, Appl
39	44	86.3	44	1	PCT-US03-12044-15	Sequence 15, Appl
40	44	86.3	44	1	PCT-US03-12044A-15	Sequence 37, Appl
41	44	86.3	51	1	PCT-US01-50250-37	Sequence 37, Appl
42	43	84.3	51	1	PCT-US01-50250A-37	Sequence 37, Appl
43	43	84.3	51	43	US-10-032-393-37	Sequence 27, Appl
44	43	84.3	80	1	PCT-US01-50250-27	
45	43	84.3				

ALIGNMENTS

RESULT 1
PCT-US01-50250-36
Sequence 36, Application PC/TUS0150250
GENERAL INFORMATION:
APPLICANT: ELITRA PHARMACEUTICALS, INC.
APPLICANT: HASELBECK, Robert
APPLICANT: WALL, Daniel
APPLICANT: GROSS, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010VPC
CURRENT APPLICATION NUMBER: PCT/US01/50250
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: US 09/948,993
PRIOR FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36

LENGTH: 51
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Xyl-T5 promoter sequence
PCT-US01-50250-36
Query Match
Best Local Similarity 100.0%; Score 51; DB 1; length 51;
Matches 51; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
Db 1 TCATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 2
PCT-US01-50250A-36
Sequence 36, Application PC/TUS0150250A
GENERAL INFORMATION:
APPLICANT: ELITRA PHARMACEUTICALS, INC.
APPLICANT: HASELBECK, Robert
APPLICANT: WALL, Daniel
APPLICANT: GROSS, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010VPC
CURRENT APPLICATION NUMBER: PCT/US01/50250A
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: 09/948,993
PRIOR FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 51
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Xyl-T5 promoter sequence
PCT-US01-50250A-36

Query Match
Best Local Similarity 100.0%; Score 51; DB 1; length 51;
Matches 51; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
Db 1 TCATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 3
US-10-032-393-36
Sequence 36, Application US/10032393
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Wall, Daniel
APPLICANT: GROSS, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010A
CURRENT APPLICATION NUMBER: US/10/032,393
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: 09/948,993
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 51
TYPE: DNA

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Xyl-T5 promoter sequence
US-10-032-393-36

Query Match 100.0%; Score 51; DB 43; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
Db

RESULT 4

PCT-US01-05269-6
Sequence 6, Application PC/TUS0105269
GENERAL INFORMATION:
APPLICANT: The Board of Trustees of the Leland Stanford Junior
TITLE OF INVENTION: Altered Recombinases for Genome Modification
FILE REFERENCE: 8400-0011.40
CURRENT APPLICATION NUMBER: PCT/US01/05269
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 72
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 1
PCT-US01-05269-6

Query Match 100.0%; Score 51; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
2 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 52
Db

RESULT 5

US-09-788-297-6
Sequence 6, Application US/09788297
GENERAL INFORMATION:
APPLICANT: Carlos, Michele P.
APPLICANT: Scilimenti, Christopher R.
TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION
FILE REFERENCE: 8400-0011
CURRENT APPLICATION NUMBER: US/09/788,297
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 72
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 1
US-09-788-297-6

Query Match 100.0%; Score 51; DB 32; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
2 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 52
Db

RESULT 6

PCT-US01-05269-7/c
Sequence 7, Application PC/TUS0105269
GENERAL INFORMATION:
APPLICANT: The Board of Trustees of the Leland Stanford Junior
TITLE OF INVENTION: Altered Recombinases for Genome Modification
FILE REFERENCE: 8400-0011.40
CURRENT APPLICATION NUMBER: PCT/US01/05269
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 7
LENGTH: 80
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 2
PCT-US01-05269-7

Query Match 100.0%; Score 51; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
75 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 25
Db

RESULT 7

PCT-US01-50250-26
Sequence 26, Application PC/TUS0150250
GENERAL INFORMATION:
APPLICANT: ELITRA PHARMACEUTICALS, INC.
APPLICANT: HASELBECK, Robert
APPLICANT: WALL, Daniel
APPLICANT: GROSS, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010VPC
CURRENT APPLICATION NUMBER: PCT/US01/50250
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: US 09/948,993
PRIOR FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 80
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Xyl-T5 fusion sequence
PCT-US01-50250-26

Query Match 100.0%; Score 51; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
1 TCATAAAAAATTATTGCTTTCAGGAAAAATTTCTGTATATAGATTCA 51
Db

RESULT 8

PCT-US01-50250A-26
Sequence 26, Application PC/TUS0150250A
GENERAL INFORMATION:
APPLICANT: ELITRA PHARMACEUTICALS, INC.
APPLICANT: HASELBECK, Robert
APPLICANT: WALL, Daniel
APPLICANT: GROSS, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010VPC


```
/ CURRENT APPLICATION NUMBER: PCT/US01/50250A
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 60/259,434
/ PRIOR FILING DATE: 2000-12-27
/ PRIOR APPLICATION NUMBER: 09/948,993
/ PRIOR FILING DATE: 2001-09-06
/ NUMBER OF SEQ ID NOS: 68
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 80
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Xyl-T5 fusion sequence
PCT-US01-50250A-26
```

```
Query Match          100.0%; Score 51; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
Db 1 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
```

RESULT 9

```
US-09-788-297-7/c
/ Sequence 7, Application US/09788297
/ GENERAL INFORMATION:
/ APPLICANT: Calos, Michele P.
/ TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION
/ FILE REFERENCE: 8400-0011
/ CURRENT APPLICATION NUMBER: US/09/788,297
/ CURRENT FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 34
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 80
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 2
US-09-788-297-7
```

```
Query Match          100.0%; Score 51; DB 32; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
Db 75 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 25
```

RESULT 10

```
US-10-032-393-26
/ Sequence 26, Application US/10032393
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Gross, Molly
/ TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
/ FILE REFERENCE: ELITRA.010A
/ CURRENT APPLICATION NUMBER: US/10/032,393
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 60/259,434
/ PRIOR FILING DATE: 2000-12-27
/ PRIOR APPLICATION NUMBER: 09/948,993
/ PRIOR FILING DATE: 2001-09-06
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ NUMBER OF SEQ ID NOS: 68
/ SOFTWARE: FastSeq for Windows Version 4.0
```

```
/ SEQ ID NO 26
/ LENGTH: 80
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Xyl-T5 fusion sequence
US-10-032-393-26
```

```
Query Match          100.0%; Score 51; DB 43; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
Db 1 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
```

RESULT 11

```
PCT-US01-50250-5
/ Sequence 5, Application PC/TUS0150250
/ GENERAL INFORMATION:
/ APPLICANT: ELITRA PHARMACEUTICALS, INC.
/ APPLICANT: HASELBECK, Robert
/ APPLICANT: WALL, Daniel
/ TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
/ FILE REFERENCE: ELITRA.010VPC
/ CURRENT APPLICATION NUMBER: PCT/US01/50250
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: US 60/259,434
/ PRIOR FILING DATE: 2000-12-27
/ PRIOR APPLICATION NUMBER: US 09/948,993
/ PRIOR FILING DATE: 2001-09-06
/ NUMBER OF SEQ ID NOS: 68
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 94
/ TYPE: DNA
/ ORGANISM: Escherichia coli
/ FEATURE:
/ OTHER INFORMATION: Oligonucleotide Xyl-T5
PCT-US01-50250-5
```

```
Query Match          100.0%; Score 51; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 51
    |||||||
Db 6 TCATATAAAATTATTGCTTTCAGGAAATTTTCTGTATAATAGATTCA 56
```

RESULT 12

```
PCT-US01-50250A-5
/ Sequence 5, Application PC/TUS0150250A
/ GENERAL INFORMATION:
/ APPLICANT: ELITRA PHARMACEUTICALS, INC.
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: WALL, Daniel
/ TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
/ FILE REFERENCE: ELITRA.010VPC
/ CURRENT APPLICATION NUMBER: PCT/US01/50250A
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 60/259,434
/ PRIOR FILING DATE: 2000-12-27
/ PRIOR APPLICATION NUMBER: 09/948,993
/ PRIOR FILING DATE: 2001-09-06
/ NUMBER OF SEQ ID NOS: 68
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 94
/ TYPE: DNA
```

ORGANISM: Escherichia coli
FEATURE:
OTHER INFORMATION: Oligonucleotide Xyl-T5
PCT-US01-50250A-5

Query Match 100.0%; Score 51; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51
Db 6 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 56

RESULT 13

US-10-032-393-5
Sequence 5, Application US/10032393
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Wall, Daniel
APPLICANT: Gross, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010A
CURRENT APPLICATION NUMBER: US/10/032,393
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: 09/948,993
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 94
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
OTHER INFORMATION: Oligonucleotide Xyl-T5
US-10-032-393-5

Query Match 100.0%; Score 51; DB 43; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51
Db 6 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 56

RESULT 14

US-60-259-434-3
Sequence 3, Application US/60259434
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010PR
CURRENT APPLICATION NUMBER: US/60/259,434
CURRENT FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 94
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial Oligonucleotide
US-60-259-434-3

Query Match 100.0%; Score 51; DB 82; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51
Db 6 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 56

RESULT 15

PCT-US01-50250-6/c
Sequence 6, Application PC/TUS0150250
GENERAL INFORMATION:

APPLICANT: ELITRA PHARMACEUTICALS, INC.
APPLICANT: HASELBECK, Robert
APPLICANT: WALL, Daniel
APPLICANT: GROSS, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010VPC
CURRENT APPLICATION NUMBER: PCT/US01/50250
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: US 09/948,993
PRIOR FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 95
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide Xyl-T5 complement
PCT-US01-50250-6

Query Match 100.0%; Score 51; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 51
Db 93 TCATAAAAAATTATTGCTTCAGGAAATTTTCTGTATAATAGATTCA 43

Search completed: May 4, 2004, 13:15:15
Job time : 3251 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 10:37:58 ; Search time 122 Seconds
(without alignments)
354.765 Million cell updates/sec

Title: US-10-032-393-36
Perfect score: 51
Sequence: 1 tcataaaaaattttatttgcgt.....ttttctgtatataatca 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1490725 seqs, 424326345 residues

Total number of hits satisfying chosen parameters: 2981450

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	27.4	53.7	111469	6	US-10-767-471-10820 Sequence 10820, A
C 2	27	52.9	4519	6	US-10-808-570-6 Sequence 6, Appli
3	26.6	52.2	7028	7	US-60-552-390-835 Sequence 835, App
4	26	51.0	201	7	US-60-563-440-35861 Sequence 35861, A
5	26	51.0	187790	7	US-60-563-440-12163 Sequence 12163, A
C 6	25.6	50.2	5777	6	US-10-821-234-51 Sequence 51, Appli
C 7	25.4	49.8	843	6	US-10-724-972A-3436 Sequence 3436, Ap
8	25.4	49.8	405740	6	US-10-796-307-8716 Sequence 8716, Ap
C 9	25	49.0	1634	1	PCT-US03-36895A-66 Sequence 66, Appli
C 10	25	49.0	40000	6	US-10-767-471-10915 Sequence 10915, A
C 11	25	49.0	57201	7	US-60-563-440-12146 Sequence 12146, A
12	25	49.0	88457	6	US-10-796-280-12211 Sequence 12211, A
13	25	49.0	130030	6	US-10-767-471-10854 Sequence 10854, A
C 14	25	49.0	193757	6	US-10-767-471-10614 Sequence 10614, A
15	25	49.0	583315	7	US-60-550-051-3041 Sequence 3041, Ap
16	25	49.0	1790242	6	US-10-767-471-10805 Sequence 10805, A
17	24.8	48.6	61635	6	US-10-765-790-105 Sequence 105, App
18	24.8	48.6	61635	6	US-10-765-790-118 Sequence 118, App
19	24.6	48.2	19810	7	US-60-563-440-12277 Sequence 12277, A
20	24.6	48.2	30510	6	US-10-796-280-12429 Sequence 12429, A
C 21	24.6	48.2	35330	6	US-10-767-471-10625 Sequence 10625, A
22	24.6	48.2	84445	6	US-10-601-886A-31 Sequence 31, Appli
23	24.6	48.2	344805	6	US-10-779-271-1 Sequence 1, Appli
24	24.6	48.2	354592	6	US-10-765-790-70 Sequence 70, Appli
C 25	24.6	48.2	1126118	6	US-10-767-471-10615 Sequence 10615, A
C 26	24.6	48.2	1126118	6	US-10-796-280-12479 Sequence 12479, A

C 27	24.6	48.2	1126118	6	US-10-796-307-8715 Sequence 8715, Ap
C 28	24.4	47.8	19053	7	US-60-563-440-12000 Sequence 12000, A
29	24.2	47.5	1650	6	US-10-797-366-254 Sequence 254, App
30	24.2	47.5	1650	6	US-10-771-187-254 Sequence 254, App
31	24.2	47.5	16201	6	US-10-767-471-10705 Sequence 10705, A
C 32	24.2	47.5	33231	6	US-10-765-790-45 Sequence 45, Appl
33	24.2	47.5	61528	7	US-60-563-440-12210 Sequence 12210, A
34	24.2	47.5	68196	6	US-10-417-375A-160 Sequence 160, App
35	24.2	47.5	68196	6	US-10-417-375B-160 Sequence 160, App
36	24	47.1	606	6	US-10-363-345A-37629 Sequence 37629, A
C 37	24	47.1	606	6	US-10-363-345A-37630 Sequence 37630, A
38	24	47.1	1200	6	US-10-767-701-13802 Sequence 13802, A
C 39	24	47.1	35126	6	US-10-796-307-8785 Sequence 8785, Ap
C 40	24	47.1	74870	6	US-10-796-280-12297 Sequence 12297, A
41	24	47.1	205784	7	US-60-550-051-3026 Sequence 3026, Ap
C 42	24	47.1	305312	7	US-60-550-051-2989 Sequence 2989, Ap
C 43	24	47.1	1790242	6	US-10-767-471-10805 Sequence 10805, A
44	23.8	46.7	201	6	US-10-767-471-16374 Sequence 16374, A
C 45	23.8	46.7	201	6	US-10-796-280-46448 Sequence 46448, A

ALIGNMENTS

RESULT 1
US-10-767-471-10820/c
Sequence 10820, Application US/10767471
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001505
CURRENT APPLICATION NUMBER: US/10/767,471
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 50231
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10820
LENGTH: 111469
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(111469)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-767-471-10820

Query Match 53.7%; Score 27.4; DB 6; Length 111469;
Best Local Similarity 75.6%; Pred. No. 30;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 7 AAAATTATTGCTTCAGGAATTTCTGTATATAGATTCA 51
DB 39669 AATTTCATTGATTTCCAAACATTTTCTGTATACATTTCA 39625

RESULT 2
US-10-808-570-6
Sequence 6, Application US/10808570
GENERAL INFORMATION:
APPLICANT: Patrick J. Dillon et al.
TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
FILE REFERENCE: PB324D1
CURRENT APPLICATION NUMBER: US/10/808,570
CURRENT FILING DATE: 2004-03-25
PRIOR APPLICATION NUMBER: US/09/956,004
PRIOR FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: 08/976,259
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/061,953
PRIOR FILING DATE: 1997-10-14
PRIOR APPLICATION NUMBER: 60/031,626
PRIOR FILING DATE: 1996-11-22
NUMBER OF SEQ ID NOS: 142

GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/563,440
; CURRENT FILING DATE: 2004-04-20
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12146
; LENGTH: 57201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-563-440-12146

Query Match 49.0%; Score 25; DB 7; Length 57201;
Best Local Similarity 75.6%; Pred. No. 1.4e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 TAAAAAATTATTGCTTCAGGAAAAATTTCTGTATAAT 44
Db 23118 TAAATGATGATGCTTCAGGAAAAATTTATATCTCTATT 23078

RESULT 12
US-10-796-280-12211
; Sequence 12211, Application US/10796280

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12211
; LENGTH: 88457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-12211

Query Match 49.0%; Score 25; DB 6; Length 88457;
Best Local Similarity 69.4%; Pred. No. 1.4e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TCATAAAAAATTATTGCTTCAGGAAAAATTTCTGTATAATAGATT 49
Db 31986 TCTTAGAAATTATTTCCTTATGCAATATGTTCAATTTCTATTGATT 32034

RESULT 13
US-10-767-471-10854
; Sequence 10854, Application US/10767471

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10854
; LENGTH: 130030
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-10854

Query Match 49.0%; Score 25; DB 6; Length 130030;

Best Local Similarity 69.4%; Pred. No. 1.4e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 3 ATAAAAAATTATTGCTTCAGGAAAAATTTCTGTATAATAGATTCA 51
Db 79661 ATATAATAATATTCTTCTTCTTATGATTCTTATAATATTTC 79709

RESULT 14
US-10-767-471-10614/c

; Sequence 10614, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10614
; LENGTH: 193757
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-10614

Query Match 49.0%; Score 25; DB 6; Length 193757;
Best Local Similarity 69.4%; Pred. No. 1.4e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 3 ATAAAAAATTATTGCTTCAGGAAAAATTTCTGTATAATAGATTCA 51
Db 27836 ATATAATAATATTCTTCTTCTTATGATTCTTATAATATTTC 27788

RESULT 15
US-60-550-051-3041

; Sequence 3041, Application US/60550051
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001508
; CURRENT APPLICATION NUMBER: US/60/550,051
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 23014
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3041
; LENGTH: 583315
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(583315)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (See Tables 1
US-60-550-051-3041

Query Match 49.0%; Score 25; DB 7; Length 583315;
Best Local Similarity 75.6%; Pred. No. 1.4e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 TAAAAAATTATTGCTTCAGGAAAAATTTCTGTATAAT 44
Db 352670 TGACAAATTATTGCTTCAGGAAAAATTTGTTCTTTATT 352710

Search completed: May 4, 2004, 13:17:32
Job time : 132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 09:47:38 ; Search time 2695 Seconds
(without alignments)
565.110 Million cell updates/sec

Title: US-10-032-393-36
Perfect score: 51
Sequence: 1 tcataaaaatttatttgc.....tttctgtataatagattca 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vr1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	30.2	59.2	318	14	CF350907	CF350907 r156h06.y
2	29.4	57.6	358	9	AW049113	AW049113 UI-M-BH1-
3	29.4	57.6	534	10	BB365884	BB365884 BB365884
4	29.4	57.6	667	10	BB080413	BB080413 BB080413

5	29.4	57.6	768	13	BQ174160	BQ174160 UI-M-DJ2-
6	29.4	57.6	2505	11	AK048039	AK048039 Mus muscu
7	29.4	57.6	3403	11	AK034193	AK034193 Mus muscu
8	29.2	57.3	356	9	AI394884	AI394884 MA001291.
9	29.2	57.3	412	9	AA219512	AA219512 zq99g04.r
10	29	56.9	130	28	B2128980	B2128980 CH230-452
11	29	56.9	595	28	B2145989	B2145989 CH230-452
12	29	56.9	1048	28	B2184761	B2184761 CH230-423
13	28.8	56.5	560	28	B2194126	B2194126 CH230-465
14	28.8	56.5	654	28	B2216912	B2216912 CH230-417
15	28.8	56.5	773	14	CF406767	CF406767 CH3#044.C
16	28.8	56.5	875	28	B2192372	B2192372 CH230-465
17	28.4	55.7	412	29	AG242333	AG242333 Lotus cor
18	28.2	55.3	140	28	AF056259	AF056259 AF056259
19	28.2	55.3	611	28	BH741163	BH741163 gt28b09.g
20	28.2	55.3	733	28	BH957079	BH957079 odj02a12.
21	28.2	55.3	757	28	BH992006	BH992006 oe111c02.
22	28.2	55.3	834	13	BU746323	BU746323 CH3#003.D
23	28	54.9	659	14	CB431012	CB431012 606949.MA
24	28	54.9	729	14	CB430297	CB430297 606181.MA
25	27.8	54.5	381	10	AW522148	AW522148 UI-R-B00-
26	27.8	54.5	748	28	BH670666	BH670666 BOMBIC18TF
27	27.6	54.1	438	28	AQ782840	AQ782840 HS.5380.B
28	27.6	54.1	824	28	BH669814	BH669814 BOMAE23TF
29	27.4	53.7	571	28	BH506288	BH506288 BOGRV16TF
30	27.4	53.7	603	9	AA799484	AA799484 EST188981
31	27.4	53.7	641	28	BZ082391	BZ082391 11e41b09.
32	27.4	53.7	702	29	CE003137	CE003137 tigr-gss-
33	27.4	53.7	718	28	BH995398	BH995398 oe929g06.
34	27.4	53.7	760	28	BH605836	BH605836 BOGXW23TR
35	27.4	53.7	993	28	BH692567	BH692567 BOMJY76TR
36	27.4	53.7	1340	29	CG751618	CG751618 P046-1-D0
37	27.2	53.3	101	28	BZ217382	BZ217382 CH230-250
38	27.2	53.3	292	9	AA365198	AA365198 EST76173
39	27.2	53.3	357	28	BZ269137	BZ269137 CH230-375
40	27.2	53.3	378	28	BZ199225	BZ199225 CH230-495
41	27.2	53.3	450	29	CE706869	CE706869 tigr-gss-
42	27.2	53.3	544	28	BZ217207	BZ217207 CH230-372
43	27.2	53.3	546	28	BZ177519	BZ177519 CH230-493
44	27.2	53.3	560	28	BZ214923	BZ214923 CH230-466
45	27.2	53.3	631	28	BZ219091	BZ219091 CH230-259

ALIGNMENTS

RESULT 1
CF350907 318 bp mRNA linear EST 20-AUG-2003
LOCUS r156h06.y1 Meloidogyne javanica J2 SMART pGEM Meloidogyne javanica
DEFINITION CDNA 5', mRNA sequence.
ACCESSION CF350907
VERSION CF350907.1 GI:33953420
KEYWORDS EST.
SOURCE Meloidogyne javanica (root-knot nematode)
ORGANISM Meloidogyne javanica
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE 1 (bases 1 to 318)
McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,B., Bennett,J., Franklin,C.,
Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999

TITLE JOURNAL
COMMENT

The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

Fax: 314 286 1810
Email: est@watson.wustl.edu
Cloned unidirectionally. Poly(A)+ RNA was concentrated and purified using Dynabeads (Dyna) and mRNA eluted for first strand synthesis. First strand cDNA was created using MMLV RT (Powerscript, Clontech) and primed with oligo(dT) with XhoI site and 5'SMART 'anchor' added using chimeric DNA-RNA oligo. 12 PCR cycles were done using first strand and primers specific to SMART oligo and 3' end. Double stranded cDNA was digested using XhoI/NotI, fractioned on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as host cells. Library materials provided by Dr. David Bird of North Carolina State University. Library construction by Jeff Rousch. See www.nematode.net for additional project information.
Putative full length read
The vector to vector length is 319
Seq primer: Sp6.

FEATURES

source

Location/Qualifiers

1..318

/organism="Meloidogyne javanica"

/mol_type="mRNA"

/db_xref="taxon:6303"

/tissue_type="whole organism"

/dev_stage="J2"

/lab_host="DH10B"

/clone_lib="Meloidogyne javanica J2 SMART pGEM"

/note="Vector: plasmid (ampicillin resistant); Site_1:

XhoI; Site_2: NotI; Cloned unidirectionally. Poly(A)+ RNA

was concentrated and purified using Dynabeads (Dyna) and

mRNA eluted for first strand synthesis. First strand cDNA

was created using MMLV RT (Powerscript, Clontech) and

primed with oligo(dT) with XhoI site and 5'SMART 'anchor'

added using chimeric DNA-RNA oligo. 12 PCR cycles were

done using first strand and primers specific to SMART

oligo and 3' end. Double stranded cDNA was digested using

XhoI/NotI, fractioned on Chroma-spin 400 columns

(Clontech) and ligated to digested pGEM-11zf(+) plasmid.

Chemically competent DH10B cells were used as host cells.

Library materials provided by Dr. David Bird of North

Carolina State University. Library construction by Jeff

Rousch. See www.nematode.net for additional project

information."

ORIGIN

Query Match

59.2%; Score 30.2; DB 14; Length 318;

Best Local Similarity 81.4%; Pred. No. 2.9e+02;

Matches 35; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTGCTTCAGGAAATTTTCTGTATAT 44

Db 80 CAAAAGAAATTATTATTATATCAGGGAATTTTCTGTATAT 122

RESULT 2
AW049113

LOCUS 358 bp mRNA linear EST 18-SEP-1999

DEFINITION UI-M-BH1-ant-g-11-0-UI.s1 NIH BMAP_M_S2 Mus musculus cDNA clone

UI-M-BH1-ant-g-11-0-UI 3', mRNA sequence.

ACCESSION AW049113

VERSION AW049113.1 GI:5909642

KEYWORDS

SOURCE EST.

ORGANISM Mus musculus (house mouse)

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 358)

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT Contact: Chin, H

National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized basal ganglia library cDNA Library Preparation: M.B.
Soares lab clone distribution: NIH BMAP cDNA clones will be made
available by the means that is soon to be determined. When NIH
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
The following repetitive elements were found in this cDNA sequence:
15-92, >(TA)n#Simple_repeat
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..358

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH1-ant-g-11-0-UI"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH BMAP_M_S2"

/note="Vector: pRTT3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The

NIH BMAP_M_S2 library is a subtracted library derived from

NIH_BMAP_M_S1, which in turn is a subtracted library

derived from a mixture of normalized libraries from ten

regions of the mouse brain (cerebellum, brain stems,

olfactory bulbs, hypothalamus, cortex, amygdala, basal

ganglia, pineal gland, striatum, hippocampus). The driver

used for subtraction consisted of a pool of 5,000 clones

from the NIH BMAP_M_S1 library and a pool of 2,000 clones

obtained from non-normalized and normalized mouse brain

spinal cord libraries.

TAG_TISSUE=basal-ganglia

TAG_LIB=NIH BMAP_M_S2

TAG_SEQ=TGTAC"

ORIGIN

Query Match

57.6%; Score 29.4; DB 9; Length 358;

Best Local Similarity 76.6%; Pred. No. 4.5e+02;

Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 ATAAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATT 49

Db 20 ATAATATATTATTGTCATACAGAAATATTATTCTATATACATT 66

RESULT 3
BB365884/c

LOCUS 534 bp mRNA linear EST 24-OCT-2001

DEFINITION BB365884 RIKEN full-length enriched, 16 days embryo head Mus

musculus cDNA clone C130031F03 3', mRNA sequence.

ACCESSION BB365884

VERSION BB365884.2 GI:16406384

KEYWORDS

SOURCE EST.

ORGANISM Mus musculus (house mouse)

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 534)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,

Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,

Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,

Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

source

1. .667
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9330162L04"
/sex="male"
/tissue_type="diencephalon"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male diencephalon"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified Bluescript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

ORIGIN

Query Match 57.6%; Score 29.4; DB 10; Length 667;
Best Local Similarity 76.6%; Pred. No. 3.7e+02;
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 3 ATAAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATT 49
|||||
Db 662 ATATATATATTATTGTCATACAGAAATATTATTCTATTATACATT 616

RESULT 5
BQ174160 768 bp mRNA linear EST 30-APR-2002
LOCUS
DEFINITION UI-M-DJ2-bvw-c-09-0-UI.s1 NIH_BMAP_DJ2 Mus musculus cDNA clone
UI-M-DJ2-bvw-c-09-0-UI 3', mRNA sequence.
BQ174160
ACCESSION BQ174160.1 GI:20349651
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 768)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
Tissue Procurement: Dr. Robin Davisson
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

FEATURES
source

The following repetitive elements were found in this cDNA
sequence: 16-93, >(TA)n#simple_repeat
Seq primer: M13 FORWARD
POLYA=Yes.
location/Qualifiers
1. .768
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-DJ2-bvw-c-09-0-UI"
/tissue_type="subfornical organ and postrema"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="NIH_BMAP DJ2"
/note="Organ: brain; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-M-DJ2 is a subtrated cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is: GCTACATGAT, subfornical organ and area postrema.
TAG_TISSUE=subfornical organ and postrema
TAG_LIB=UI-M-DJ2
TAG_SEQ=GCTACATGAT"

ORIGIN

Query Match 57.6%; Score 29.4; DB 13; Length 768;
Best Local Similarity 76.6%; Pred. No. 3.5e+02;
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 3 ATAAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATT 49
|||||
Db 21 ATATATATATTATTGTCATACAGAAATATTATTCTATTATACATT 67

RESULT 6
AK048039/c 2505 bp mRNA linear HTC 20-SEP-2003
LOCUS
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130031F03 product:unknown EST, full insert
sequence.
AK048039
ACCESSION AK048039.1 GI:26092610
VERSION
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

```

TITLE
JOURNAL MEDLINE PUBMED
REFERENCE AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2505)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoch,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES
source
1..2505
Location/Qualifiers
1..2505
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C130031F03"
/db_xref="MGI:2414069"
/db_xref="taxon:10090"
/clone="C130031F03"
/tissue_type="head"
/clone_1db="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
1..2505
/note="unknown EST (GB|B1735511, evidence: BLASTN, 99%,
match=817)"

misc_feature

ORIGIN
Query Match 57.6%; Score 29.4; DB 11; Length 2505;
Best Local Similarity 76.6%; Pred. No. 2.3e+02;
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 3 ATAAATAATTTATTTCCTTCAGGAATTTTCCTGTATATAGATT 49
||||||| ||||| ||||| ||||| ||||| ||||| |||||
2274 ATAATATATTTATTGGCATACAGAAATATTATTTCTATTAATCAATT 2228
```

RESULT 7	
AK034193/c	
LOCUS	3403 bp mRNA linear HTC 18-SEP-2003
DEFINITION	Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330162L04 product:unknown EST, full insert sequence.
ACCESSION	AK034193
VERSION	AK034193.1 GI:26083816
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3403)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katch, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,

Fax:81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.
URL: <http://fantom.gsc.riken.go.jp/>.
Location/Qualifiers

FEATURES
source
1. 3403
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9330162L04"
/db_xref="MGI:2358388"
/db_xref="taxon:10090"
/clone="9330162L04"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. 3403
/note="unknown EST (GB|BI735511, evidence: BLASTN, 99%, match=817)"

ORIGIN
Query Match 57.6%; Score 29.4; DB 11; Length 3403;
Best Local Similarity 76.6%; Pred. No. 2.1e+02;
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

CY 3 ATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATT 49
|||||
Db 3391 ATATATATTTATTTGCATACAGAAATATTTCTATATACATT 3345
|||||

RESULT 8 356 bp mRNA linear EST 04-FEB-1999
LOCUS MA001291.C8F Soares normalized S8W Schistosoma mansoni cDNA 3',
DEFINITION mRNA sequence.
ACCESSION AI394884
VERSION AI394884
KEYWORDS GI:4224431
SOURCE EST.
ORGANISM Schistosoma mansoni
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoida; Schistosomatidae; Schistosoma.
1 (bases 1 to 356)
Bailey, J.A., Bentley, K., Blanton, R.E., Soares, M.B. and Chakravarti, A.
Expressed sequence tags from a normalized adult Schistosoma mansoni library
Unpublished (1999)
JOURNAL Contact: Chakravarti A
COMMENT Department of Genetics
Case Western Reserve University
2109 Adelbert Rd, Cleveland, OH 44106, USA
Tel: 216-368-5847
Fax: 216-368-5857
Email: axc39@po.cwru.edu
Additional data regarding this EST may be found at
<http://genome.cwru.edu/schistosoma/est/S8West.html>
Seq primer: M13 Forward.
Features
Location/Qualifiers
1. 356
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/strain="Puerto Rican"
/db_xref="taxon:6183"
/sex="mixed"
/dev_stage="8 week old adult worms"
/lab_host="Mus musculus"

/clone_lib="Soares normalized S8W"
/note="Vector: pT7T3-pac; Site_1: NotI; Site_2: EcoRI;
Note: normalization has resulted in an increased frequency of the rare contaminating mouse (host) sequences (~10%).
For details of library construction see: Bonaldo MF, Lennon G, Soares MB. Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. 1996. Genome Research 6:791-806"

ORIGIN
Query Match 57.3%; Score 29.2; DB 9; Length 356;
Best Local Similarity 74.0%; Pred. No. 5.1e+02;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

CY 2 CATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
|||||
Db 140 CATTAATAATCTAATCTTCTTAACAATAATCTTCTGGAATAGATTAA 189
|||||

RESULT 9 412 bp mRNA linear EST 07-FEB-1997
LOCUS AA219512/c
DEFINITION zq99g04.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
CDNA clone IMAGE:650166 5', mRNA sequence.
ACCESSION AA219512
VERSION AA219512.1 GI:1833571
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 412)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 262.
Features
Location/Qualifiers
1. 412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5276755"
/db_xref="taxon:9606"
/clone="IMAGE:650166"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/note="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

ORIGIN
Query Match 57.3%; Score 29.2; DB 9; Length 412;

Best Local Similarity 72.5%; Pred. No. 4.8e+02;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCTGTATATAGATCA 51
|||||
122 TCNTAAAAAGATGATTTCTTCTTAATAAATTTCTGTATCTTAATTTA 72

RESULT 10
BZ128980 130 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-452M22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-452M22, genomic survey sequence.
ACCESSION BZ128980
VERSION BZ128980.1 GI:23769927
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 130)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 452 row: M column: 22
Seq primer: SP6
Class: BAC ends.

FEATURES
Location/Qualifiers
source 1..130
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNhsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-452M22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SSNhsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
1..130
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNhsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-452M22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SSNhsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 56.9%; Score 29; DB 28; Length 130;
Best Local Similarity 86.5%; Pred. No. 8e+02;
Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 AAAAAATTTATTTGCTTTCAGGAAATTTTCTGTAT 41
|||||
10 AAAAAATTTGATTCATTCAGAAATTTTCTGTAT 46

RESULT 11
BZ145989 595 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-452N22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-452N22, genomic survey sequence.
ACCESSION BZ145989

VERSION BZ145989.1 GI:23786936
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 595)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-452N22.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 452 row: N column: 22
Seq primer: SP6
Class: BAC ends.

FEATURES
Location/Qualifiers
source 1..595
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNhsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-452N22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SSNhsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
1..595
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNhsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-452N22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SSNhsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 56.9%; Score 29; DB 28; Length 595;
Best Local Similarity 86.5%; Pred. No. 4.8e+02;
Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 AAAAAATTTATTTGCTTTCAGGAAATTTTCTGTAT 41
|||||
Db 485 AAAAAATTTGATTCATTCAGAAATTTTCTGTAT 521

RESULT 12
BZ184761 1048 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-423F16.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-423F16, genomic survey sequence.
ACCESSION BZ184761
VERSION BZ184761.1 GI:23834700
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1048)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)

COMMENT Other_GSSs: CH230-423F16.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 423 row: F column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES Location/Qualifiers
1..1048
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-423F16"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 56.9%; Score 29; DB 28; Length 1048;
Best Local Similarity 86.5%; Pred. No. 4e+02;
Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 AAAAAATTATTGCTTTCAGGAAATTTTCTGTAT 41
|||||
108 AAAAAATTGATTCATTCATGAATTTTCTGTAT 72

Db

RESULT 13
BZ194126/c 560 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-465D15.TJB CHORI-230 Segment 2 Rattus norvegicus genomic
DEFINITION clone CH230-465D15, genomic survey sequence.
BZ194126
ACCESSION BZ194126
VERSION BZ194126.1 GI:23852178
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 560)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other_GSSs: CH230-465D15.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_ering_information.htm). BAC end
(http://www.chori.org/bacpac/or_ering_information.htm). BAC end

TITLE
JOURNAL
COMMENT

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 465 row: D column: 15
Seq primer: SP6
Class: BAC ends.

FEATURES Location/Qualifiers
1..560
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-465D15"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 56.5%; Score 28.8; DB 28; Length 560;
Best Local Similarity 82.5%; Pred. No. 5.5e+02;
Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTGCTTTCAGGAAATTTTCTGTAT 41
|||||
Db 526 CCTAGAAATTTAATTCATTCATGAATTTTCTGTAT 487

RESULT 14
BZ216912/c 654 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-41703.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-41703, genomic survey sequence.
BZ216912
ACCESSION BZ216912
VERSION BZ216912.1 GI:23875270
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 654)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other_GSSs: CH230-41703.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 417 row: O column: 3
Seq primer: SP6
Class: BAC ends.

FEATURES Location/Qualifiers
1..654
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-41703"
/sex="Female"
/cell_type="Brain"

TITLE
JOURNAL
COMMENT

Job time : 2708 secs

/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARBA1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SeNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 56.5%; Score 28.8; DB 28; Length 654;
Best Local Similarity 82.5%; Pred. No. 5.2e+02;
Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTTGCTTTCAGGAAATTTTCTGTAT 41
| | | | | | | | | | | | | | | | | | | | | |
Db 513 CTTAGAAATTTAATTCATTCATGAAATTTTCTGTAT 474

RESULT 15

CF406767/c 773 bp mRNA linear EST 02-SEP-2003
LOCUS CH3#044_C08T3 Canine heart normalized cDNA library in pBluescript
DEFINITION Canis familiaris cDNA clone CH3#044_C08 3', mRNA sequence.

ACCESSION CF406767
VERSION CF406767.1 GI:34407011
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
TITLE 1 (bases 1 to 773).
JOURNAL Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.
COMMENT Expressed sequence tags from Canine heart
Unpublished (2003)
Other_ESTs: CH3#044_C08T7
Contact: George AL
Division of Genetic Medicine
Vanderbilt University
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
Tel: 615 936 2660
Fax: 615 936 2661

Email: al.george@vanderbilt.edu
Insert Length: 1448 Std Error: 0.00
Seq primer: T3: ATTAACCTCCTCACTAAAGGA
High quality sequence start: 64
High quality sequence stop: 733.

FEATURES

source
1..773
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CH3#044_C08"
/tissue_type="heart"
/cell_type="heart"
/dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
/clone_lib="Canine heart normalized cDNA library in pBluescript"
/note="Organ: heart; Vector: pBluescript; Site_1: 5' of vector NotI; Site_2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

ORIGIN

Query Match 56.5%; Score 28.8; DB 14; Length 773;
Best Local Similarity 75.0%; Pred. No. 4.9e+02;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTTGCTTTCAGGAAATTTTCTGTATAATAGATT 49
| | | | | | | | | | | | | | | | | | | | | |
Db 238 CTTAAAAAATTTTAGTGATAGAAATTAATTTTCTATATGATAGACT 191